

(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 747 483 A2

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:

11.12.1996 Bulletin 1996/50

(51) Int. Cl.<sup>6</sup>: C12N 15/53, C12N 15/54,  
C12N 15/60, C12N 15/70,  
C12N 15/74, C12N 15/80,  
C12N 15/81, C12P 23/00

(21) Application number: 96108556.0

(22) Date of filing: 29.05.1996

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB IT LI NL• Pasamontes, Luis  
4632 Trimbach (CH)

(30) Priority: 09.06.1995 EP 95108888

• Tessier, Michel  
68100 Mulhouse (FR)(71) Applicant: F. HOFFMANN-LA ROCHE AG  
4002 Basel (CH)• van Loon, Adolphus  
4310 Rheinfelden (CH)

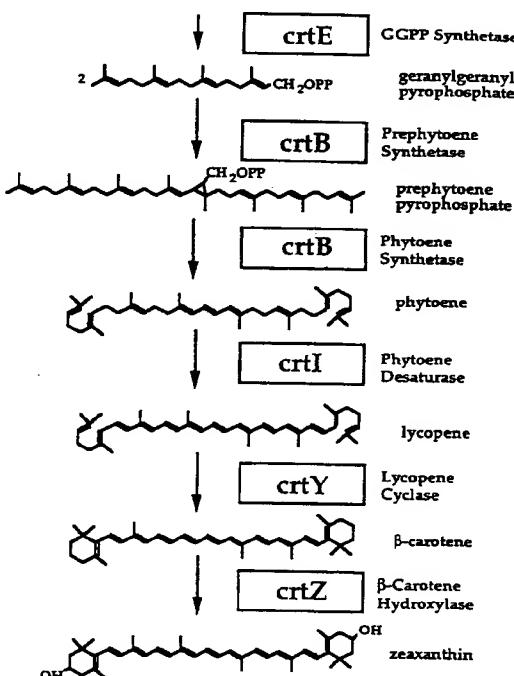
(72) Inventors:

• Hohmann, Hans-Peter  
79100 Freiburg (DE)(74) Representative: Braun, Axel et al  
F.Hoffmann-La Roche AG  
Patent Department (PLP),  
124 Grenzacherstrasse  
4070 Basel (CH)

## (54) Fermentative carotenoid production

(57) The present invention is directed to a DNA sequence comprising one or more DNA sequences selected from the group consisting of a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE), a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB), a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI), a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which encodes the  $\beta$ -carotene hydroxylase of *Flavobacterium* sp. R1534 (crtZ) or DNA sequences which are substantially homologous, vectors comprising such DNA sequences and/or a DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of *Alcaligenes* strain PC-1 (crtW) or a DNA sequence which is substantially homologous, cells which are transformed by such DNA sequences and/or vectors, a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such transformed cells and a process for the preparation of a food or feed composition.

Fig. 1



EP 0 747 483 A2

## Description

Over 600 different carotenoids have been described from carotenogenic organisms found among bacteria, yeast, fungi and plants. Currently only two of them,  $\beta$ -carotene and astaxanthin are commercially produced in microorganisms and used in the food and feed industry.  $\beta$ -carotene is obtained from algae and astaxanthin is produced in *Pfaffia* strains which have been generated by classical mutation. However, fermentation in *Pfaffia* has the disadvantage of long fermentation cycles and recovery from algae is cumbersome. Therefore it is desirable to develop production systems which have better industrial applicability, e.g. can be manipulated for increased titers and/or reduced fermentation times. Two such systems using the biosynthetic genes from *Erwinia herbicola* and *Erwinia uredovora* have already been described in WO 91/13078 and EP 393 690, respectively. Furthermore, three  $\beta$ -carotene ketolase genes ( $\beta$ -carotene  $\beta$ -4-oxygenase) of the marine bacteria *Agrobacterium aurantiacum* and *Alcaligenes* strain PC-1 (crtW) [Misawa, 1995, Biochem. Biophys. Res. Com. 209, 867-876][Misawa, 1995, J. Bacteriology 177, 6575-6584] and from the green algae *Haematococcus pluvialis* (bkt) [Lotan, 1995, FEBS Letters 364, 125-128][Kajiwara, 1995, Plant Mol. Biol. 29, 343-352] have been cloned. *E. coli* carrying either the carotenogenic genes (crtE, crtB, crtY and crtI) of *E. herbicola* [Hundle, 1994, MGG 245, 406-416] or of *E. uredovora* and complemented with the crtW gene of *A. aurantiacum* [Misawa, 1995] or the bkt gene of *H. pluvialis* [Lotan, 1995][Kajiwara, 1995] resulted in the accumulation of canthaxanthin ( $\beta,\beta$ -carotene-4,4'-dione), originating from the conversion of  $\beta$ -carotene, via the intermediate echinenone ( $\beta,\beta$ -carotene-4-one). Introduction of the above mentioned genes (crtW or bkt) into *E. coli* cells harbouring besides the carotenoid biosynthesis genes mentioned above also the crtZ gene of *E. uredovora* [Kajiwara, 1995][Misawa, 1995], resulted in both cases in the accumulation of astaxanthin (3,3'-dihydroxy- $\beta,\beta$ -carotene-4,4'-dione). The results obtained with the bkt gene, are in contrast to the observation made by others [Lotan, 1995], who using the same experimental set-up, but introducing the *H. pluvialis* bkt gene in a zeaxanthin ( $\beta,\beta$ -carotene-3,3'-diol) synthesising *E. coli* host harbouring the carotenoid biosynthesis genes of *E. herbicola*, a close relative of the above mentioned *E. uredovora* strain, did not observe astaxanthin production.

However, functionally active combinations of the carotenoid biosynthesising genes of the present invention with the known crtW genes have not been shown so far and even more importantly there is a continuing need in even more optimized fermentation systems for industrial application.

It is therefore an object of the present invention to provide a DNA sequence comprising one or more DNA sequences selected from the group consisting of:

- 30 a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous;
- 35 b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) or a DNA sequence which is substantially homologous;
- 40 c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI) or a DNA sequence which is substantially homologous;
- 45 d) a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which is substantially homologous;
- 50 e) a DNA sequence which encodes the  $\beta$ -carotene hydroxylase of *Flavobacterium* sp. R1534 (crtZ) or a DNA sequence which is substantially homologous.

It is also an object of the present invention to provide a vector comprising such DNA sequence, preferably in the form of an expression vector. Furthermore it is an object of the present invention to provide a cell which is transformed by such DNA sequence or vector, preferably which is a prokaryotic cell and more preferably which is *E. coli* or a *Bacillus* strain. Such transformed cell which is an eukaryotic cell, preferably a yeast cell or a fungal cell, is also an object of the present invention. Finally the present invention concerns also a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such a cell under suitable culture conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells or the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid or carotenoid mixture is added to food or feed.

Furthermore, a DNA sequence comprising the following DNA sequences is an object of the present invention:

- a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous, and

b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) of a DNA sequence which is substantially homologous, and

5 c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R 1534 (crtI) or a DNA sequence which is substantially homologous.

It is also an object of the present invention to provide a vector comprising such DNA sequence, preferably in the form of an expression vector. Furthermore it is an object of the present invention to provide a cell which is transformed by such DNA sequence or vector, preferably which is a prokaryotic cell and more preferably which is *E. coli* or a *Bacillus* strain. Such transformed cell which is an eukaryotic cell, preferably a yeast cell or a fungal cell, is also an object of the present invention. Finally the present invention concerns also a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such a cell under suitable conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells or the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present, preferably such a process for the preparation of lycopene and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid, preferably lycopene or carotenoid mixture, preferably a lycopene comprising mixture is added to food or feed.

Furthermore a DNA sequence comprising the following DNA sequence is also an object of the present invention:

20 a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous, and

b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) or a DNA sequence which is substantially homologous, and

25 c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI) or a DNA sequence which is substantially homologous, and

d) a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which is substantially homologous.

It is also an object of the present invention to provide a vector comprising such DNA sequence, preferably in the form of an expression vector. Furthermore it is an object of the present invention to provide a cell which is transformed by such DNA sequence or vector, preferably which is a prokaryotic cell and more preferably which is *E. coli* or a *Bacillus* strain. Such transformed cell which is an eukaryotic cell, preferably a yeast cell or a fungal cell, is also an object of the present invention. Finally the present invention concerns also a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such a cell under suitable conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells or the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present, preferably such a process for the preparation of  $\beta$ -carotene and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid, preferably  $\beta$ -carotene or carotenoid mixture, preferably a  $\beta$ -carotene comprising mixture is added to food or feed.

Furthermore a cell which is transformed by the above mentioned DNA sequence comprising subsequences a) to d) or the vector comprising it and a second DNA sequence which encodes the  $\beta$ -carotene  $\beta 4$ -oxygenase of *Alcaligenes* strain PC-1 (crt W) or a DNA sequence which is substantially homologous or a second vector which comprises a DNA sequence which encodes the  $\beta$ -carotene  $\beta 4$ -oxygenase of *Alcaligenes* strain PC-1 (crt W) or a DNA sequence which is substantially homologous; and a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such cell under suitable conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells or the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present, preferably such a process for the preparation of echinenone and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid, preferably echinenone or carotenoid mixture, preferably an echinenone comprising mixture is added to food or feed.

Furthermore it is an object of the present invention to provide a DNA sequence as mentioned above comprising subsequences a) to d) and a DNA sequence which encodes the  $\beta$ -carotene  $\beta 4$ -oxygenase of *Alcaligenes* strain PC-1 (crt W) or a DNA sequence which is substantially homologous and a vector comprising such DNA sequence, preferably in the form of an expression vector. Furthermore it is an object of the present invention to provide a cell which is transformed by such DNA sequence or vector, preferably which is a prokaryotic cell and more preferably which is *E. coli* or a *Bacillus* strain. Such transformed cell which is an eukaryotic cell, preferably a yeast cell or a fungal cell, is also an

object of the present invention. Finally the present invention concerns also a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such a cell under suitable culture conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells of the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present, especially such a process for the preparation of echinenone or canthaxanthin and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid, preferably echinenone or canthaxanthin or carotenoid mixture, preferably a echinenone or canthaxanthin containing mixture is added to food or feed.

Furthermore a DNA sequence comprising the following DNA sequences is also an object of the present invention:

- 10 a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous, and
- 15 b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) or a DNA sequence which is substantially homologous, and
- 20 c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtl) or a DNA sequence which is substantially homologous, and
- 25 d) a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which is substantially homologous, and
- 30 e) a DNA sequence which encodes the  $\beta$ -carotene hydroxylase of *Flavobacterium* sp. R1534 (crtZ) or a DNA sequence which is substantially homologous.

It is also an object of the present invention to provide a vector comprising such DNA sequence, preferably in the form of an expression vector. Furthermore it is an object of the present invention to provide a cell which is transformed by such DNA sequence or vector, preferably which is a prokaryotic cell and more preferably which is *E. coli* or a *Bacillus* strain. Such transformed cell which is an eukaryotic cell, preferably a yeast cell or a fungal cell, is also an object of the present invention. Finally the present invention concerns also a process for the preparation of a desired carotenoid or mixture of carotenoids by culturing such a cell under suitable conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells of the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present, preferably such a process for the preparation of zeaxanthin and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid, preferably zeaxanthin or the carotenoid mixture, preferably a zeaxanthin containing mixture is added to food or feed.

Furthermore a DNA sequence as mentioned above comprising subsequences a) to e) and in addition a DNA sequence which encodes the  $\beta$ -carotene  $\beta 4$ -oxygenase of *Alcaligenes* strain PC-1 (crt W) or a DNA sequence which is substantially homologous is an object of the present invention and to provide a vector comprising such DNA sequence, preferably in form of an expression vector. Furthermore it is an object of the present invention to provide a cell which is transformed by such DNA sequence or vector, preferably which is a prokaryotic cell and more preferably which is *E. coli* or a *Bacillus* strain. Such transformed cell which is an eukaryotic cell, preferably a yeast cell or a fungal cell, is also an object of the present invention. Finally the present invention concerns also a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such a cell under suitable culture conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells of the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present, preferably such a process for the preparation of zeaxanthin, adonixanthin or astaxanthin and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid, preferably zeaxanthin, adonixanthin or astaxanthin or carotenoid mixture, preferably a zeaxanthin, adonixanthin or astaxanthin containing mixture is added to food or feed.

Furthermore a cell which is transformed by the DNA sequence mentioned above comprising subsequences a) to e) or a vector comprising such DNA sequence and a second DNA sequence which encodes the  $\beta$ -carotene  $\beta 4$ -oxygenase of *Alcaligenes* strain PC-1 (crt W) or a DNA sequence which is substantially homologous or a second vector which comprises a DNA sequence which encodes the  $\beta$ -carotene  $\beta 4$ -oxygenase of *Alcaligenes* strain PC-1 (crt W) or a DNA sequence which is substantially homologous is also an object of the present invention and a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such a cell under suitable conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells or the culture medium, and in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present, preferably such a process for the preparation of zeaxanthin or adonixanthin and a process for the preparation of a food or feed composition characterized therein that after such a process has been effected the carotenoid, preferably zeaxanthin or

adonixanthin or carotenoid mixture, preferably a zeaxanthin or adonixanthin containing mixture is added to food or feed.

In this context it should be mentioned that the expression "a DNA sequence is substantially homologous" refers with respect to the crtE encoding DNA sequence to a DNA sequence which encodes an amino acid sequence which shows more than 45 %, preferably more than 60 % and more preferably more than 75 % and most preferably more than 90 % identical amino acids when compared to the amino acid sequence of crtE of *Flavobacterium* sp. 1534 and is the amino acid sequence of a polypeptide which shows the same type of enzymatic activity as the enzyme encoded by crtE of *Flavobacterium* sp. 1534. In analogy with respect to crtB this means more than 60 %, preferably more than 70 %, more preferably more than 80 % and most preferably more than 90 %; with respect to crtI this means more than 70 %, preferably more than 80 % and most preferably more than 90 %; with respect to crtY this means 55 %, preferably 70 %, more preferably 80 % and most preferably 90 %; with respect to crtZ this means more than 60 %, preferably 70 %, more preferably 80 % and most preferably 90 %; with respect to crtW this also means more than 60 %, preferably 70 %, more preferably 80 % and most preferably 90 %. Sequences which are substantially homologous to crtW are known, e.g. in form of the  $\beta$ -carotene  $\beta 4$ -oxygenase of *Agrobacterium aurantiacum* or the green algae *Haematococcus pluvialis* (bkt).

DNA sequences in form of genomic DNA, cDNA or synthetic DNA can be prepared as known in the art [see e.g.

15 Sambrook et al., *Molecular Cloning*, Cold Spring Harbor Laboratory Press 1989] or, e.g. as specifically described in Examples 1, 2 or 7.

The cloning of the DNA-sequences of the present invention from such genomic DNA can than be effected, e.g. by using the well known polymerase chain reaction (PCR) method. The principles of this method are outlined e.g. in *PCR Protocols: A guide to Methods and Applications*, Academic Press, Inc. (1990). PCR is an in vitro method for producing 20 large amounts of a specific DNA of defined length and sequence from a mixture of different DNA-sequences. Thereby, PCR is based on the enzymatic amplification of the specific DNA fragment of interest which is flanked by two oligonucleotide primers which are specific for this sequence and which hybridize to the opposite strand of the target sequence. The primers are oriented with their 3' ends pointing toward each other. Repeated cycles of heat denaturation of the template, annealing of the primers to their complementary sequences and extension of the annealed primers with a DNA 25 polymerase result in the amplification of the segment between the PCR primers. Since the extension product of each primer can serve as a template for the other, each cycle essentially doubles the amount of the DNA fragment produced in the previous cycle. By utilizing the thermostable Taq DNA polymerase, isolated from the thermophilic bacteria *Ther-30 mus aquaticus*, it has been possible to avoid denaturation of the polymerase which necessitated the addition of enzyme after each heat denaturation step. This development has led to the automation of PCR by a variety of simple tempera-35 ture-cycling devices. In addition, the specificity of the amplification reaction is increased by allowing the use of higher temperatures for primer annealing and extension. The increased specificity improves the overall yield of amplified products by minimizing the competition by non-target fragments for enzyme and primers. In this way the specific sequence of interest is highly amplified and can be easily separated from the non-specific sequences by methods known in the art, e.g. by separation on an agarose gel and cloned by methods known in the art using vectors as described e.g. by Holten and Graham in *Nucleic Acid Res.* 19, 1156 (1991), Kovalic et. al. in *Nucleic Acid Res.* 19, 4560 (1991), Marchuk et al. in *Nucleic Acid Res.* 19, 1154 (1991) or Mead et al. in *Bio/Technology* 9, 657-663 (1991).

The oligonucleotide primers used in the PCR procedure can be prepared as known in the art and described e.g. in Sambrook et al., s.a.

40 Amplified DNA-sequences can than be used to screen DNA libraries by methods known in the art (Sambrook et al., s.a.) or as specifically described in Examples 1 and 2.

Once complete DNA-sequences of the present invention have been obtained they can be used as a guideline to define new PCR primers for the cloning of substantially homologous DNA sequences from other sources. In addition they and such homologous DNA sequences can be integrated into vectors by methods known in the art and described e.g. in Sambrook et al. (s.a.) to express or overexpress the encoded polypeptide(s) in appropriate host systems. However, a man skilled in the art knows that also the DNA-sequences themselves can be used to transform the suitable host 45 systems of the invention to get overexpression of the encoded polypeptide. Appropriate host systems are for example Bacteria e.g. *E. coli*, *Bacilli* as, e.g. *Bacillus subtilis* or *Flavobacter* strains. *E. coli*, which could be used are *E. coli* K12 strains e.g. M15 [described as DZ 291 by Villarejo et al. in *J. Bacteriol.* 120, 466-474 (1974)], HB 101 [ATCC No. 33694] or *E. coli* SG13009 [Gottesman et al., *J. Bacteriol.* 148, 265-273 (1981)]. Suitable eukaryotic host systems are for example fungi, like *Aspergilli*, e.g. *Aspergillus niger* [ATCC 9142] or yeasts, like *Saccharomyces cerevisiae* or *Pichia*, like *pastoris*, all available from ATCC.

50 Suitable vectors which can be used for expression in *E. coli* are mentioned, e.g. by Sambrook et al. [s.a.] or by Fiers et al. in *Proc. 8th Int. Biotechnology Symposium* [Soc. Franc. de Microbiol., Paris (Durand et al., eds.), pp. 680-697 (1988)] or by Bujard et al. in *Methods in Enzymology*, eds. Wu and Grossmann, Academic Press, Inc. Vol. 155, 416-433 (1987) and Stüber et al. in *Immunological Methods*, eds. Lefkovits and Pernis, Academic Press, Inc. Vol. IV, 121-152 (1990). Vectors which could be used for expression in *Bacilli* are known in the art and described, e.g. in EP 405 370, EP 635 572 *Proc. Nat. Acad. Sci. USA* 81, 439 (1984) by Yansura and Henner, *Meth. Enzym.* 185, 199-228 (1990) or EP 207 459. Vectors which can be used for expression in fungi are known in the art and described e.g. in EP 420 358 and for yeast in EP 183 070, EP 183 071, EP 248 227, EP 263 311.

Once such DNA-sequences have been expressed in an appropriate host cell in a suitable medium the carotenoids can be isolated either from the medium in the case they are secreted into the medium or from the host organism and, if necessary separated from other carotenoids if present in case one specific carotenoid is desired by methods known in the art (see e.g. Carotenoids Vol IA: Isolation and Analysis, G. Britton, S. Liaaen-Jensen, H. Pfander; 1995, Birkhäuser Verlag, Basel).

The carotenoids of the present invention can be used in a process for the preparation of food or feeds. A man skilled in the art is familiar with such process. Such compound foods or feeds can further comprise additives or components generally used for such purpose and known in the state of the art.

After the invention has been described in general hereinbefore, the following figures and examples are intended to illustrate details of the invention, without thereby limiting it in any matter.

Figure 1: The biosynthesis pathway for the formation of carotenoids of *Flavobacterium* sp. R1534 is illustrated explaining the enzymatic activities which are encoded by DNA sequences of the present invention.

Figure 2: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested with the restriction enzymes shown on top of each lane and hybridized with Probe 46F. The arrow indicated the isolated 2.4 kb XbaI/PstI fragment.

Figure 3: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested with CiaI or double digested with CiaI and HindIII. Blots shown in Panel A and B were hybridized to probe A or probe B, respectively (see examples). Both CiaI/HindIII fragments of 1.8 kb and 9.2 kb are indicated.

Figure 4: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested with the restriction enzymes shown on top of each lane and hybridized to probe C. The isolated 2.8 kb SacI/HindIII fragment is shown by the arrow.

Figure 5: Southern blot of genomic *Flavobacterium* sp. R1534 DNA digested with the restriction enzymes shown on top of each lane and hybridized to probe D. The isolated BclI/SphI fragment of approx. 3 kb is shown by the arrow.

Figure 6: Physical map of the organization of the carotenoid biosynthesis cluster in *Flavobacterium* sp. R1534, deduced from the genomic clones obtained. The location of the probes used for the screening are shown as bars on the respective clones.

Figure 7: Nucleotide sequence of the *Flavobacterium* sp. R1534 carotenoid biosynthesis cluster and its flanking regions. The nucleotide sequence is numbered from the first nucleotide shown (see BamHI site of Fig. 6). The deduced amino acid sequence of the ORF's (orf-5, orf-1, crtE, crtB, crtI, crtY, crtZ and orf-16) are shown with the single-letter amino acid code. Arrow (--) indicate the direction of the transcription; asterisks, stop codons.

Figure 8: Protein sequence of the GGPP synthase (crtE) of *Flavobacterium* sp. R1534 with a MW of 31331 Da.

Figure 9: Protein sequence of the prephytoene synthetase (crtB) of *Flavobacterium* sp. R1534 with a MW of 32615 Da.

Figure 10: Protein sequence of the phytoene desaturase (crtI) of *Flavobacterium* sp. R1534 with a MW of 54411 Da.

Figure 11: Protein sequence of the lycopene cyclase (crtY) of *Flavobacterium* sp. R1534 with a MW of 42368 Da.

Figure 12: Protein sequence of the  $\beta$ -carotene hydroxylase (crtZ) of *Flavobacterium* sp. R1534 with a MW of 19282 Da.

Figure 13: Recombinant plasmids containing deletions of the *Flavobacterium* sp. R1534 carotenoid biosynthesis gene cluster.

Figure 14: Primers used for PCR reactions. The underlined sequence is the recognition site of the indicated restriction enzyme. Small caps indicate nucleotides introduced by mutagenesis. Boxes show the artificial RBS which is recognized in *B. subtilis*. Small caps in bold show the location of the original adenine creating

the translation start site (ATG) of the following gene (see original operon). All the ATG's of the original Flavobacter carotenoid biosynthetic genes had to be destroyed to not interfere with the rebuild transcription start site. Arrows indicate start and ends of the indicated *Flavobacterium R1534 WT* carotenoid genes.

5      **Figure 15:** Linkers used for the different constructions. The underlined sequence is the recognition site of the indicated restriction enzyme. Small caps indicate nucleotides introduced by synthetic primers. Boxes show the artificial RBS which is recognized in *B. subtilis*. Arrow indicate start and ends of the indicated *Flavobacterium* carotenoid genes.

10     **Figure 16:** Construction of plasmids pBIKS(+)-clone59-2, pLyco and pZea4.

15     **Figure 17:** Construction of plasmid p602CAR.

20     **Figure 18:** Construction of plasmids pBIKS(+)-CARVEG-E and p602 CARVEG-E.

25     **Figure 19:** Construction of plasmids pHP13-2CARZYB-EINV and pHP13-2PN25ZYB-EINV.

30     **Figure 20:** Construction of plasmid pXI12-ZYIB-EINVMUTRBS2C.

35     **Figure 21:** Northern blot analysis of *B. subtilis* strain BS1012::ZYIB-EINV4. Panel A: Schematic representation of a reciprocal integration of plasmid pXI12-ZYIB-EINV4 into the levan-sucrase gene of *B. subtilis*. Panel B: Northern blot obtained with probe A (PCR fragment which was obtained with CAR 51 and CAR 76 and hybridizes to the 3' end of crtZ and the 5' end of crtY). Panel C: Northern blot obtained with probe B (BamHI-Xhol fragment isolated from plasmid pBIKS(+)-crtE/2 and hybridizing to the 5' part of the crtE gene).

40     **Figure 22:** Schematic representation of the integration sites of three transformed *Bacillus subtilis* strains: BS1012::SFCO, BS1012::SFCOCAT1 and BA1012::SFCONEO1. Amplification of the synthetic Flavobacterium carotenoid operon (SFCO) can only be obtained in those strains having amplifiable structures. Probe A was used to determine the copy number of the integrated SFCO. Erythromycine resistance gene (ermAM), chloramphenicol resistance gene (cat), neomycine resistance gene (neo), terminator of the cryT gene of *B. subtilis* (cryT), levan-sucrase gene (sac-B 5' and sac-B 3'), plasmid sequences of pXI12 (pXI12), promoter originating from site I of the veg promoter complex (Pvegl).

45     **Figure 23:** Construction of plasmids pXI12-ZYIB-EINV4MUTRBS2CNEO and pXI12-ZYIB-EINV4MUTRBS2CCAT.

50     **Figure 24:** Complete nucleotide sequence of plasmid pZea4.

55     **Figure 25:** Synthetic crtW gene of *Alcaligenes PC-1*. The translated protein sequence is shown above the double stranded DNA sequence. The twelve oligonucleotides (crtW1-crtW12) used for the PCR synthesis are underlined.

60     **Figure 26:** Construction of plasmid pBIKS-crtEBIYZW. The HindIII-Pm11 fragment of pALTER-Ex2-crtW, carrying the synthetic crtW gene, was cloned into the HindIII and MluI (blunt) sites. Pvegl and Ptac are the promoters used for the transcription of the two operas. The ColE1 replication origin of this plasmid is compatible with the p15A origin present in the pALTER-Ex2 constructs.

65     **Figure 27:** Relevant inserts of all plasmids constructed in Example 7. Disrupted genes are shown by //. Restriction sites: S=SacI, X=XbaI, H=HindIII, N=NsiI, Hp=HpaI, Nd=NdeI.

70     **Figure 28:** Reaction products (carotenoids) obtained from  $\beta$ -carotene by the process of the present invention.

75     **Example 1**

#### Materials and general methods used

**Bacterial strains and plasmids:** *Flavobacterium sp. R1534 WT* (ATCC 21588) was the DNA source for the genes

cloned. Partial genomic libraries of *Flavobacterium sp.* R1534 WT DNA were constructed into the pBluescriptII+(KS) or (SK) vector (Stratagene, La Jolla, USA) and transformed into *E. coli* XL-1 blue (Stratagene) or JM109.

5 **Media and growth conditions:** Transformed *E. coli* were grown in Luria broth (LB) at 37° C with 100µg Ampicillin (Amp)/ml for selection. *Flavobacterium sp.* R1534 WT was grown at 27° C in medium containing 1% glucose, 1% tryptone (Difco Laboratories), 1% yeast extract (Difco), 0.5% MgSO<sub>4</sub> 7H<sub>2</sub>O and 3% NaCl.

**Colony screening:** Screening of the *E. coli* transformants was done by PCR basically according to the method described by Zon et al. [Zon et al., BioTechniques 7, 696-698 (1989)] using the following primers:

10 Primer #7: 5'-CCTGGATGACGTGCTGGAATATTCC-3'

Primer #8: 5'-CAAGGCCAGATCGCAGGCG-3'

15 **Genomic DNA:** A 50 ml overnight culture of *Flavobacterium sp.* R1534 was centrifuged at 10,000 g for 10 minutes. The pellet was washed briefly with 10 ml of lysis buffer (50 mM EDTA, 0.1M NaCl pH7.5), resuspended in 4 ml of the same buffer supplemented with 10 mg of lysozyme and incubated at 37°C for 15 minutes. After addition of 0.3 ml of N-Lauroyl sarcosine (20%) the incubation at 37°C was continued for another 15 minutes before the extraction of the DNA with phenol, phenol/chloroform and chloroform. The DNA was ethanol precipitated at room temperature for 20 minutes in the presence of 0.3 M sodium acetate (pH 5.2), followed by centrifugation at 10,000 g for 15 minutes. The pellet was rinsed with 70% ethanol, dried and resuspended in 1 ml of TE (10 mM Tris, 1mM EDTA, pH 8.0).

20 All genomic DNA used in the southern blot analysis and cloning experiments was dialysed against H<sub>2</sub>O for 48 hours, using collodium bags (Sartorius, Germany), ethanol precipitated in the presence of 0.3 M sodium acetate and resuspended in H<sub>2</sub>O.

25 **Probe labelling:** DNA probes were labeled with ( $\alpha$ -<sup>32</sup>P) dGTP (Amersham) by random-priming according to [Sambrook et al., s.a.].

30 **Probes used to screen the mini-libraries:** **Probe 46F** is a 119 bp fragment obtained by PCR using primer #7 and #8 and *Flavobacterium sp.* R1534 genomic DNA as template. This probe was proposed to be a fragment of the *Flavobacterium sp.* R1534 phytoene synthase (crtB) gene, since it shows significant homology to the phytoene synthase genes from other species (e.g. *E. uredovora*, *E. herbicola*). **Probe A** is a BstXI - PstI fragment of 184 bp originating from the right arm of the insert of clone 85. **Probe B** is a 397 bp Xhol - NotI fragment obtained from the left end of the insert of clone 85. **Probe C** is a 536 bp BglII - PstI fragment from the right end of the insert of clone 85. **Probe D** is a 376 bp KpnI - BstYI fragment isolated from the insert of clone 59. The localization of the individual probes is shown in figure 6.

35 **Oligonucleotide synthesis:** The oligonucleotides used for PCR reactions or for sequencing were synthesized with an Applied Biosystems 392 DNA synthesizer.

40 **Southern blot analysis:** For hybridization experiments *Flavobacterium sp.* R1534 genomic DNA (3 µg) was digested with the appropriate restriction enzymes and electrophoresed on a 0.75% agarose gel. The transfer to Zeta-Probe blotting membranes (BIO-RAD), was done as described [Southern, E.M., J. Mol. Biol. 98, 503 (1975)]. Prehybridization and hybridization was in 7%SDS, 1% BSA (fraction V; Boehringer), 0.5M Na<sub>2</sub>HPO<sub>4</sub>, pH 7.2 at 65°C. After hybridization the membranes were washed twice for 5 minutes in 2x SSC, 1% SDS at room temperature and twice for 15 minutes in 0.1% SSC, 0.1% SDS at 65° C.

45 **DNA sequence analysis:** The sequence was determined by the dideoxy chain termination technique [Sanger et al., Proc. Natl. Acad. Sci. USA 74, 5463-5467 (1977)] using the Sequenase Kit (United States Biochemical). Both strands were completely sequenced and the sequence analyzed using the GCG sequence analysis software package (Version 8.0) by Genetics Computer, Inc. [Devereux et al., Nucleic Acids. Res. 12, 387-395 (1984)].

50 **Analysis of carotenoids:** *E. coli* XL-1 or JM109 cells (200 - 400 ml) carrying different plasmid constructs were grown for the times indicated in the text, usually 24 to 60 hours, in LB supplemented with 100µg Ampicillin/ml, in shake flasks at 37° C and 220 rpm.

55 The carotenoids present in the microorganisms were extracted with an adequate volume of acetone using a rotation homogenizer (Polytron, Kinematica AG, CH-Luzern). The homogenate was filtered through the sintered glass of a suction filter into a round bottom flask. The filtrate was evaporated by means of a rotation evaporator at 50° C using a water-jet vacuum. For the zeaxanthin detection the residue was dissolved in n-hexane/acetone (86:14) before analysis with a normalphase HPLC as described in [Weber, S. in Analytical Methods for Vitamins and Carotenoids in Feed, Keller, H.E. , Editor, 83-85 (1988)]. For the detection of  $\beta$ -carotene and lycopene the evaporated extract was dissolved in n-hexane/acetone (99:1) and analysed by HPLC as described in [Hengartner et al., Helv. Chim. Acta 75, 1848-1865 (1992)].

Example 2Cloning of the *Flavobacterium sp. R1534* carotenoid biosynthetic genes.

5 To identify and isolate DNA fragments carrying the genes of the carotenoid biosynthesis pathway, we used the DNA fragment 46F (see methods) to probe a Southern Blot carrying chromosomal DNA of *Flavobacterium sp. R1534* digested with different restriction enzymes Fig. 2. The 2.4 kb Xhol/PstI fragment hybridizing to the probe seemed the most appropriate one to start with. Genomic *Flavobacterium sp. R1534* DNA was digested with Xhol/PstI and run on a 1% agarose gel. According to a comigrating DNA marker, the region of about 2.4 kb was cut out of the gel and the DNA 10 isolated. A Xhol/PstI mini library of *Flavobacterium sp. R1534* genomic DNA was constructed into Xhol - PstI sites of pBluescriptIISK(+). One hundred *E. coli* XL1 transformants were subsequently screened by PCR with primer #7 and primer # 8, the same primers previously used to obtain the 119 bp fragment (46F). One positive transformant, named done 85, was found. Sequencing of the insert revealed sequences not only homologous to the phytoene synthase (crtB) but also to the phytoene desaturase (crtl) of both *Erwinia* species *herbicola* and *uredovora*. Left and right hand 15 genomic sequences of done 85 were obtained by the same approach using probe A and probe B. *Flavobacterium sp. R1534* genomic DNA was double digested with ClaI and Hind III and subjected to Southern analysis with probe A and probe B. With probe A a ClaI/HindIII fragment of approx. 1.8 kb was identified (Fig. 3A), isolated and subcloned into the ClaI/HindIII sites of pBluescriptIISK (+). Screening of the *E. coli* XL1 transformants with probe A gave 6 positive clones. The insert of one of these positives, clone 43-3, was sequenced and showed homology to the N-terminus of crtI genes. 20 and to the C terminus of crtY genes of both *Erwinia* species mentioned above. With probe B an approx. 9.2 kb ClaI/HindIII fragment was detected (Fig. 3B), isolated and subcloned into pBluescriptIISK (+).

A screening of the transformants gave one positive, clone 51. Sequencing of the 5' and 3' of the insert, revealed 25 that only the region close to the HindIII site showed relevant homology to genes of the carotenoid biosynthesis of the *Erwinia* species mentioned above (e.g. crtB gene and crtE gene). The sequence around the ClaI site showed no homology to known genes of the carotenoid biosynthesis pathway. Based on this information and to facilitate further sequencing and construction work, the 4.2 kb BamHI/HindIII fragment of done 51 was subcloned into the respective sites of pBluescriptIISK(+) resulting in clone 2. Sequencing of the insert of this clone confirmed the presence of genes homologous to *Erwinia* sp. crtB and crtE genes. These genes were located within 1.8 kb from the HindIII site. The remaining 2.4 kb of this insert had no homology to known carotenoid biosynthesis genes.

30 Additional genomic sequences downstream of the ClaI site were detected using probe C to hybridize to *Flavobacterium sp. R1534* genomic DNA digested with different restriction enzymes (see figure 4).

A SalI/HindIII fragment of 2.8 kb identified by Southern analysis was isolated and subcloned into the HindIII/Xhol sites of pBluescriptIISK (+). Screening of the *E. coli* XL1 transformants with probe A gave one positive clone named 35 clone 59. The insert of this clone confirmed the sequence of clone 43-3 and contained in addition sequences homologous to the N-terminus of the crtY gene from other known lycopene cyclases. To obtain the putative missing crtZ gene a Sau3AI partial digestion library of *Flavobacterium sp. R1534* was constructed into the BamHI site of pBluescriptIISK(+). Screening of this library with probe D gave several positive clones. One transformant designated, clone 6a, had an insert of 4.9 kb. Sequencing of the insert revealed besides the already known sequences coding for crtB, crtI and crtY also the missing crtZ gene. Clone 7g was isolated from a mini library carrying BclI/SphI fragments of R1534 (Fig. 40 5) and screened with probe D. The insert size of done 7g is approx. 3 kb.

The six independent inserts of the clones described above covering approx. 14 kb of the *Flavobacterium sp. R1534* genome are compiled in Figure 6.

The determined sequence spanning from the BamHI site (position 1) to base pair 8625 is shown figure 7.

45 **Putative protein coding regions of the cloned R1534 sequence.**

Computer analysis using the CodonPreference program of the GCG package, which recognizes protein coding regions by virtue of the similarity of their codon usage to a given codon frequency table, revealed eight open reading frames (ORFs) encoding putative proteins: a partial ORF from 1 to 1165 (ORF-5) coding for a polypeptide larger than 50 41382 Da; an ORF coding for a polypeptide with a molecular weight of 40081 Da from 1180 to 2352 (ORF-1); an ORF coding for a polypeptide with a molecular weight of 31331 Da from 2521 to 3405 (crtE); an ORF coding for a polypeptide with a molecular weight of 32615 Da from 4316 to 3408 (crtB); an ORF coding for a polypeptide with a molecular weight of 54411 Da from 5797 to 4316 (crtl); an ORF coding for a polypeptide with a molecular weight of 42368 Da from 6942 to 5797 (crtY); an ORF coding for a polypeptide with a molecular weight of 19282 Da from 7448 to 6942 (crtZ); and an ORF coding for a polypeptide with a molecular weight of 19368 Da from 8315 to 7770 (ORF-16); ORF-1 and crtE have the opposite transcriptional orientation from the others (Fig. 6). The translation start sites of the ORFs crtI, crtY and crtZ could clearly be determined based on the appropriately located sequences homologous to the Shine/Dalgarno (S/D) [Shine and Dalgarno, Proc. Natl. Acad. Sci. USA 71, 1342-1346 (1974)] consensus sequence AGG-6-9N--ATG (Fig. 10) and the homology to the N-terminal sequences of the respective enzymes of *E. herbicola* and *E. uredovora*. The

5 translation of the ORF crtB could potentially start from three closely spaced codons ATG (4316), ATG (4241) and ATG (4211). The first one, although not having the best S/D sequence of the three, gives a translation product with the highest homology to the N-terminus of the *E. herbicola* and *E. uredovora* crtB protein, and is therefore the most likely translation start site. The translation of ORF crtE could potentially start from five different start codons found within 150 bp : ATG (2389), ATG (2446), ATG (2473), ATG (2497) and ATG (2521). We believe that based on the following observations, the ATG (2521) is the most likely transcription start site of crtE: this ATG start codon is preceeded by the best consensus S/D sequence of all five putative start sites mentioned; and the putative N-terminal amino acid sequence of the protein encoded has the highest homology to the N-terminus of the crtE enzymes of *E. herbicola* and *E. uredovora*;

## 10 Characteristics of the crt translational initiation sites and gene products.

The translational start sites of the five carotenoid biosynthesis genes are shown below and the possible ribosome binding sites are underlined. The genes crtZ, crtY, crtI and crtB are grouped so tightly that the TGA stop codon of the anterior gene overlaps the ATG of the following gene. Only three of the five genes (crtI, crtY and crtZ) fit with the consensus for optimal S/D sequences. The boxed TGA sequence shows the stop codon of the anterior gene.

20	-10	+1	
	ACGA <u>AGGCACCGATGACGCCA</u>		crtE
	<u>CGGACCTGGCCGTCGCATGACCGATC</u>		crtB
25	<u>CGGATCGCAA TACATGAGCCATG</u>		crtY
	<u>CTGCA<u>GGAGAGAGCA</u>TGAGTTCCG</u>		crtI
30	<u>GCAAGGGGCCGGCATGAGCACTT</u>		crtZ

35 Amino acid sequences of individual crt genes of *Flavobacterium* sp. R1534.

All five ORFs of *Flavobacterium* sp. R1534 having homology to known carotenoid biosynthesis genes of other species are clustered in approx. 5.2 kb of the sequence (Fig. 7).

#### GGDP synthase (*crtE*)

The amino acid (aa) sequence of the geranylgeranyl pyrophosphate synthase (crtE gene product) consists of 295 aa and is shown in figure 8. This enzyme condenses farnesyl pyrophosphate and isopentenyl pyrophosphate in a 1'-4.

### Phytoene synthase (*crtB*)

This enzyme catalyzes two enzymatic steps. First it condenses in a head to head reaction two geranylgeranyl pyrophosphates (C20) to the C40 carotenoid prephytoene. Second it rearranges the cyclopropylring of prephytoene to phytoene. The 303 aa encoded by the crtB gene of *Flavobacterium* sp. R1534 is shown in figure 9.

### ***Phytoene desaturase (crtI)***

55 The phytoene desaturase of *Flavobacterium sp.* R1534 consisting of 494 aa, shown in figure 10, performs like the  
 ctrl enzyme of *E. herbicola* and *E. uredovora*, four desaturation steps, converting the non-coloured carotenoid phy-  
 toene to the red coloured lycopene.

***Lycopene cyclase (crtY)***

The crtY gene product of *Flavobacterium sp.* R1534 is sufficient to introduce the  $\beta$ -ionone rings at both sides of lycopene to obtain  $\beta$ -carotene. The lycopene cyclase of *Flavobacterium sp.* R1534 consists of 382 aa (Fig. 11).

5

 ***$\beta$ -carotene hydroxylase (crtZ)***

The gene product of crtZ consisting of 169 aa (Fig. 12) and hydroxylates  $\beta$ -carotene to the xanthophyll zeaxanthin.

10 **Putative enzymatic functions of the ORF's (orf-1, orf-5 and orf-16)**

The orf-1 has at the aa level over 40% identity to acetoacetyl-CoA thiolases of different organisms (e.g. *Candida tropicalis*, human, rat). This gene is therefore most likely a putative acetoacetyl-CoA thiolase (acetyl-CoA acetyltransferase), which condenses two molecules of acetyl-CoA to Acetoacetyl-CoA. Condensation of acetoacetyl-CoA with a third acetyl-CoA by the HMG-CoA synthase forms  $\beta$ -hydroxy- $\beta$ -methylglutaryl-CoA (HMG-CoA). This compound is part of the mevalonate pathway which produces besides sterols also numerous kinds of isoprenoids with diverse cellular functions. In bacteria and plants, the isoprenoid pathway is also able to synthesize some unique products like carotenoids, growth regulators (e.g. in plants gibberellins and abscisic acid) and secondary metabolites like phytoalexins [Riou et al., Gene 148, 293-297 (1994)].

20 The orf-5 has a low homology of approx. 30%, to the amino acid sequence of polyketide synthases from different streptomycetes (e.g. *S. violaceoruber*, *S. cinnamonensis*). These antibiotic synthesizing enzymes (polyketide synthases), have been classified into two groups. Type-I polyketide synthases are large multifunctional proteins, whereas type-II polyketide synthases are multiprotein complexes composed of several individual proteins involved in the subreactions of the polyketide synthesis [Bibb, et al. Gene 142, 31-39 (1994)].

25 The putative protein encoded by the orf-16 has at the aa level an identity of 42% when compared to the soluble hydrogenase subunit of *Anabaena cylindrica*.

**Functional assignment of the ORF's (crtE, crtB, crtI, crtY and crtZ) to enzymatic activities of the carotenoid biosynthesis pathway.**

30

The biochemical assignment of the gene products of the different ORF's were revealed by analyzing carotenoid accumulation in *E. coli* host strains that were transformed with deleted variants of the *Flavobacterium sp.* gene cluster and thus expressed not all of the crt genes (Fig. 13).

35 Three different plasmid were constructed: pLyco, p59-2 and pZea4. Plasmid p59-2 was obtained by subcloning the HindIII/BamHI fragment of clone 2 into the HindIII/BamHI sites of clone 59. p59-2 carries the ORF's of the crtE, crtB, crtI and crtY gene and should lead to the production of  $\beta$ -carotene. pLyco was obtained by deleting the KpnI/KpnI fragment, coding for approx. one half (N-terminus) of the crtY gene, from the p59-2 plasmid. *E. coli* cells transformed with pLyco, and therefore having a truncated non-functional crtY gene, should produce lycopene, the precursor of  $\beta$ -carotene. pZea4 was constructed by ligation of the Ascl-Spel fragment of p59-2, containing the crtE, crtB, crtI and most of the crtY gene with the Ascl/XbaI fragment of clone 6a, containing the sequences to complete the crtY gene and the crtZ gene. pZea4 [for complete sequence see Fig. 24; nucleotides 1 to 683 result from pBluescriptIIKS(+), nucleotides 684 to 8961 from *Flavobacterium* R1534 WT genome, nucleotides 8962 to 11233 from pBluescriptIIKS(+)] has therefore all five ORF's of the zeaxanthin biosynthesis pathway. Plasmid pZea4 has been deposited on May 25, 1995 at the DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Germany) under accession No. DSM 10012. *E. coli* cells transformed with this latter plasmid should therefore produce zeaxanthin. For the detection of the carotenoid produced, transformants were grown for 48 hours in shake flasks and then subjected to carotenoid analysis as described in the methods section. Figure 13 summarizes the different inserts of the plasmids described above, and the main carotenoid detected in the cells.

40 45 As expected the pLyco carrying *E. coli* cells produced lycopene, those carrying p59-2 produced  $\beta$ -carotene (all-E,9-Z,13-Z) and the cells having the pZea4 construct produced zeaxanthin. This confirms that all the necessary genes of *Flavobacterium sp.* R1534 for the synthesis of zeaxanthin or their precursors (phytoene, lycopene and  $\beta$ -carotene) were cloned.

**Example 3**

55

**Materials and methods used for expression of carotenoid synthesizing enzymes**

**Bacterial strains and plasmids:** The vectors pBluescriptIIKS (+) or (-) (Stratagene, La Jolla, USA) and pUC18 [Vieira and Messing, Gene 19, 259-268 (1982); Norrander et al., Gene 26, 101-106 (1983)] were used for cloning in dif-

ferent *E. coli* strains, like XL-1 blue (Stratagene), TG1 or JM109. In all *B. subtilis* transformations, strain 1012 was used. Plasmids pHP13 [Haima et al., Mol. Gen. Genet. 209, 335-342 (1987)] and p602/22 [LeGrice, S.F.J. in Gene Expression Technology, Goeddel, D.V., Editor, 201-214 (1990)] are Gram (+)/(-) shuttle vectors able to replicate in *B. subtilis* and *E. coli* cells. Plasmid p205 contains the vegI promoter cloned into the SmaI site of pUC18. Plasmid pXI12 is an integration vector for the constitutive expression of genes in *B. subtilis* [Haiker et al., in 7th Int. Symposium on the Genetics of Industrial Microorganisms, June 26-July 1, 1994. Montreal, Quebec, Canada (1994)]. Plasmid pBEST501 [Itaya et al., Nucleic Acids Res. 17 (11), 4410 (1989)] contains the neomycin resistance gene cassette originating from the plasmid pUB110 (GenBank entry: M19465) of *S. aureus* [McKenzie et al., Plasmid 15, 93-103 (1986); McKenzie et al., Plasmid 17, 83-84 (1987)]. This neomycin gene has been shown to work as a selection marker when present in a single copy in the genome of *B. subtilis*. Plasmid pC194 (ATCC 37034)(GenBank entry: L08860) originates from *S. aureus* [Horinouchi and Weisblum, J. Bacteriol. 150, 815-825 (1982)] and contains the chloramphenicol acetyltransferase gene.

**Media and growth conditions:** *E. coli* were grown in Luria broth (LB) at 37° C with 100µg Ampicillin (Amp)/ml for selection. *B. subtilis* cells were grown in VY-medium supplemented with either erythromycin (1 µg/ml), neomycin (5-180 µg/ml) or chloramphenicol (10-80 µg/ml).

**Transformation:** *E. coli* transformations were done by electroporation using the Gen-pulser device of BIO-RAD (Hercules, CA, USA) with the following parameters (200 Ω, 250 µFD, 2.5V). *B. subtilis* transformations were done basically according to the standard procedure method 2.8 described by [Cutting and Vander Horn in Molecular Biological Methods for Bacillus, Harwood, C.R and Cutting, S.M., Editor, John Wiley & Sons: Chichester, England. 61-74 (1990)].

**Colony screening:** Bacterial colony screening was done as described by [Zon et al., s.a.].

**Oligonucleotide synthesis:** The oligonucleotides used for PCR reactions or for sequencing were synthesized with an Applied Biosystems 392 DNA synthesizer.

**PCR reactions:** The PCR reactions were performed using either the **UITma DNA polymerase** (Perkin Elmer Cetus) or the **Pfu Vent polymerase** (New England Biolabs) according to the manufacturers instructions. A typical 50 µl PCR reaction contained: 100ng of template DNA, 10 pM of each of the primers, all four dNTP's (final conc. 300 µM), MgCl<sub>2</sub> (when UITma polymerase was used; final conc. 2 mM), 1x UITma reaction buffer or 1x Pfu buffer (supplied by the manufacturer). All components of the reaction with the exception of the DNA polymerase were incubated at 95°C for 2 min. followed by the cycles indicated in the respective section (see below). In all reactions a hot start was made, by adding the polymerase in the first round of the cycle during the 72°C elongation step. At the end of the PCR reaction an aliquot was analysed on 1% agarose gel, before extracting once with phenol/chloroform. The amplified fragment in the aqueous phase was precipitated with 1/10 of a 3M NaAcetate solution and two volumes of Ethanol. After centrifugation for 5 min. at 12000 rpm, the pellet was resuspended in an adequate volume of H<sub>2</sub>O, typically 40 µl, before digestion with the indicated restriction enzymes was performed. After the digestion the mixture was separated on a 1% low melting point agarose. The PCR product of the expected size were excised from the agarose and purified using the glass beads method (GENECLEAN KIT, Bio 101, Vista CA, USA) when the fragments were above 400 bp or directly spun out of the gel when the fragments were shorter than 400 bp, as described by [Heery et al., TIBS 6 (6), 173 (1990)].

#### Oligos used for gene amplification and site directed mutagenesis:

All PCR reactions performed to allow the construction of the different plasmids are described below. All the primers used are summarized in figure 14.

Primers #100 and #101 were used in a PCR reaction to amplify the complete crtE gene having a SpeI restriction site and an artificial ribosomal binding site (RBS) upstream of the transcription start site of this gene. At the 3' end of the amplified fragment, two unique restriction sites were introduced, an AvrII and a SmaI site, to facilitate the further cloning steps. The PCR reaction was done with UITma polymerase using the following conditions for the amplification: 5 cycles with the profile: 95°C, 1 min./ 60°C, 45 sec./ 72°C, 1 min. and 20 cycles with the profile: 95°C, 1 min./ 72°C, 1 min.. Plasmid pBIKS(+)-clone2 served as template DNA. The final PCR product was digested with SpeI and SmaI and isolated using the GENECLEAN KIT. The size of the fragment was approx. 910 bp.

Primers #104 and #105 were used in a PCR reaction to amplify the crtZ gene from the translation start till the Sall restriction site, located in the coding sequence of this gene. At the 5' end of the crtZ gene an EcoRI, a synthetic RBS and a Ndel site was introduced. The PCR conditions were as described above. Plasmid pBIKS(+)-clone 6a served as template DNA and the final PCR product was digested with EcoRI and Sall. Isolation of the fragment of approx. 480 bp was done with the GENECLEAN KIT.

Primers **MUT1** and **MUT5** were used to amplify the complete crtY gene. At the 5' end, the last 23 nucleotides of the crtZ gene including the Sall site are present, followed by an artificial RBS preceding the translation start site of the crtY gene. The artificial RBS created includes a PmII restriction site. The 3' end of the amplified fragment contains 22 nucleotides of the crtY gene, preceded by a newly created artificial RBS which contains a MunI restriction site. The conditions used for the PCR reaction were as described above using the following cycling profile: 5 rounds of 95°C, 45 sec./ 60°C, 45 sec./ 72°C, 75 sec. followed by 22 cycles with the profile: 95°C, 45 sec./ 66°C, 45 sec./ 72°C, 75 sec.. Plasmid

pXI12-ZYIB-EINV4 served as template for the Pfu Vent polymerase. The PCR product of 1225 bp was made blunt and cloned into the SmaI site of pUC18, using the Sure-Clone Kit (Pharmacia) according to the manufacturer.

5 Primers **MUT2** and **MUT6** were used to amplify the complete *crtl* gene. At the 5' the last 23 nucleotides of the *crtY* gene are present, followed by an artificial RBS which precedes the translation start site of the *crtl* gene. The new RBS created, includes a *Mun*I restriction site. The 3' end of the amplified fragment contains the artificial RBS upstream of the *crtB* gene including a *Bam*HI restriction site. The conditions used for the PCR reaction were basically as described above including the following cycling profile: 5 rounds of 95°C, 30 sec./ 60°C, 30 sec./ 72°C, 75 sec., followed by 25 cycles with the profile: 95°C, 30 sec./ 66°C, 30 sec./ 72°C, 75 sec.. Plasmid pXI12-ZYIB-EINV4 served as template for the Pfu Vent polymerase. For the further cloning steps the PCR product of 1541 bp was digested with *Mun*I and *Bam*HI.

10 Primers **MUT3** and **CAR17** were used to amplify the N-terminus of the *crtB* gene. At the 5' the last 28 nucleotides of the *crtl* gene are present followed by an artificial RBS, preceding the translation start site of the *crtB* gene. This new created RBS, includes a *Bam*HI restriction site. The amplified fragment, named PCR-F contains also the *Hind*III restriction site located at the N-terminus of the *crtB* gene. The conditions used for the PCR reaction were as described elsewhere in the text, including the following cycling profile: 5 rounds of 95°C, 30 sec./ 58°C, 30 sec./ 72°C, 20 sec. followed by 25 cycles with the profile: 95°C, 30 sec./ 60°C, 30 sec./ 72°C, 20 sec.. Plasmid pXI12-ZYIB-EINV4 served as template for the Pfu Vent polymerase. The PCR product of approx. 160 bp was digested with *Bam*HI and *Hind*III.

**Oligos used to amplify the chloramphenicol resistance gene (cat):**

20 Primers **CAT3** and **CAT4** were used to amplify the chloramphenicol resistance gene of pC194 (ATCC 37034) [Horinouchi and Weisblum, s.a.] a R-plasmid found in *S. aureus*. The conditions used for the PCR reaction were as described previously including the following cycling profile: 5 rounds of 95°C, 60 sec./ 50°C, 60 sec./ 72°C, 2 min. followed by 20 cycles with the profile: 95°C, 60 sec./ 60°C, 60 sec./ 72°C, 2 min.. Plasmid pC198 served as template for the Pfu Vent polymerase. The PCR product of approx. 1050 bp was digested with *Eco*RI and *Aat*II.

25 **Oligos used to generate linkers:** Linkers were obtained by adding 90 ng of each of the two corresponding primers into an Eppendorf tube. The mixture was dried in a speed vac and the pellet resuspended in 1x Ligation buffer (Boehringer, Mannheim, Germany). The solution was incubated at 50°C for 3 min. before cooling down to RT, to allow the primers to hybridize properly. The linker were now ready to be ligated into the appropriate sites. All the oligos used to generate linkers are shown in figure 15.

30 Primers **CS1** and **CS2** were used to form a linker containing the following restrictions sites *Hind*III, *Af*III, *Sc*al, *Xba*l, *Pme*l and *Eco*RI.

Primers **MUT7** and **MUT8** were used to form a linker containing the restriction sites *Sall*, *Avr*II, *Pm*II, *Mlu*l, *Mun*l, *Bam*HI, *Sph*I and *Hind*III.

Primers **MUT9** and **MUT10** were used to introduce an artificial RBS upstream of *crtY*.

35 Primers **MUT11** and **MUT12** were used to introduce an artificial RBS upstream of *crtE*.

**Isolation of RNA:** Total RNA was prepared from log phase growing *B. subtilis* according to the method described by [Maes and Messens, Nucleic Acids Res. 20 (16), 4374 (1992)].

40 **Northern Blot analysis:** For hybridization experiments up to 30 µg of *B. subtilis* RNA was electrophoresed on a 1% agarose gel made up in 1x MOPS and 0.66 M formaldehyde. Transfer to Zeta-Probe blotting membranes (BIO-RAD), UV cross-linking, pre-hybridization and hybridization was done as described elsewhere in [Farrell, J.R.E., RNA Methodologies. A laboratory Guide for isolation and characterization. San Diego, USA: Academic Press (1993)]. The washing conditions used were: 2 x 20 min. in 2xSSPE/0.1% SDS followed by 1 x 20min. in 0.1% SSPE/0.1% SDS at 65°C. Northern blots were then analyzed either by a Phosphorimager (Molecular Dynamics) or by autoradiography on X-ray films from Kodak.

45 **Isolation of genomic DNA:** *B. subtilis* genomic DNA was isolated from 25 ml overnight cultures according to the standard procedure method 2.6 described by [13].

50 **Southern blot analysis:** For hybridization experiments *B. subtilis* genomic DNA (3 µg) was digested with the appropriate restriction enzymes and electrophoresed on a 0.75% agarose gel. The transfer to Zeta-Probe blotting membranes (BIO-RAD), was done as described [Southern, E.M., s.a.]. Prehybridization and hybridization was in 7%SDS, 1% BSA (fraction V; Boehringer), 0.5M  $\text{Na}_2\text{HPO}_4$ , pH 7.2 at 65°C. After hybridization the membranes were washed twice for 5 min. in 2x SSC, 1% SDS at room temperature and twice for 15 min. in 0.1% SSC, 0.1% SDS at 65°C. Southern blots were then analyzed either by a Phosphorimager (Molecular Dynamics) or by autoradiography on X-ray films from Kodak.

55 **DNA sequence analysis:** The sequence was determined by the dideoxy chain termination technique [Sanger et al., s.a.] using the Sequenase Kit Version 1.0 (United States Biochemical). Sequence analysis were done using the GCG sequence analysis software package (Version 8.0) by Genetics Computer, Inc. [Devereux et al., s.a.].

**Gene amplification in *B. subtilis*:** To amplify the copy number of the SFCO in *B. subtilis* transformants, a single colony was inoculated in 15 ml VY-medium supplemented with 1.5 % glucose and 0.02 mg chloramphenicol or neomycin/ml, dependend on the antibiotic resistance gene present in the amplifiable structure (see results and discussion).

The next day 750 µl of this culture were used to inoculate 13 ml VY-medium containing 1.5% glucose supplemented with (60, 80, 120 and 150 µg/ml) for the cat resistant mutants, or 160 µg/ml and 180 µg/ml for the neomycin resistant mutants). The cultures were grown overnight and the next day 50 µl of different dilutions (1: 20, 1:400, 1: 8000, 1: 160'000) were plated on VY agar plates with the appropriate antibiotic concentration. Large single colonies were then further analyzed to determine the number of copies and the amount of carotenoids produced.

5      **Analysis of carotenoids:** *E. coli* or *B. subtilis* transformants (200 - 400 ml) were grown for the times indicated in the text, usually 24 to 72 hours, in LB-medium or VY-medium, respectively, supplemented with antibiotics, in shake flasks at 37° C and 220 rpm.

10     The carotenoids produced by the microorganisms were extracted with an adequate volume of acetone using a rotation homogenizer (Polytron, Kinematica AG, CH-Luzern). The homogenate was filtered through the sintered glass of a suction filter into a round bottom flask. The filtrate was evaporated by means of a rotation evaporator at 50° C using a water-jet vacuum. For the zeaxanthin detection the residue was dissolved in n-hexane/acetone (86:14) before analysis with a normalphase HPLC as described in [Weber, S., s.a.]. For the detection of β-carotene and lycopene the evaporated extract was dissolved in n-hexane/acetone (99:1) and analysed by HPLC as described in Hengartner et al., s.a.].

15

#### Example 4

##### Carotenoid production in *E. coli*

20     The biochemical assignment of the gene products of the different open reading frames (ORF'S) of the carotenoid biosynthesis cluster of *Flavobacterium sp.* were revealed by analyzing the carotenoid accumulation in *E. coli* host strains, transformed with plasmids carrying deletions of the *Flavobacterium sp.* gene cluster, and thus lacking some of the crt gene products. Similar functional assays in *E. coli* have been described by other authors [Misawa et al., s.a.; Perry et al., J. Bacteriol., 168, 607-612 (1986); Hundle, et al., Molecular and General Genetics 254 (4), 406-416 (1994)].

25     Three different plasmid pLyco, pBIIKS(+)-clone59-2 and pZea4 were constructed from the three genomic isolates pBIIKS(+)-clone2, pBIIKS(+)-clone59 and pBIIKS(+)-clone6a (see figure 16).

Plasmid pBIIKS(+)-clone59-2 was obtained by subcloning the HindIII/BamHI fragment of pBIIKS(+)-clone 2 into the HindIII/BamHI sites of pBIIKS(+)-clone59. The resulting plasmid pBIIKS(+)-clone59-2 carries the complete ORF's of the crtE, crtB, crtI and crtY gene and should lead to the production of β-carotene. pLyco was obtained by deleting the 30 KpnI/KpnI fragment, coding for approx. one half (N-terminus) of the crtY gene, from the plasmid pBIIKS(+)-clone59-2. *E. coli* cells transformed with pLyco, and therefore having a truncated non-functional crtY gene, should produce lycopene, the precursor of β-carotene. pZea4 was constructed by ligation of the Ascl-Spel fragment of pBIIKS(+)-clone59-2, containing the crtE, crtB, crtI and most of the crtY gene with the Ascl/XbaI fragment of clone 6a, containing the crtZ gene and sequences to complete the truncated crtY gene mentioned above. pZea4 has therefore all five ORF's of the 35 zeaxanthin biosynthesis pathway. *E. coli* cells transformed with this latter plasmid should therefore produce zeaxanthin. For the detection of the carotenoid produced, transformants were grown for 43 hours in shake flasks and then subjected to carotenoid analysis as described in the methods section. Figure 16 summarizes the construction of the plasmids described above.

40     As expected the pLyco carrying *E. coli* cells produced lycopene, those carrying pBIIKS(+)-clone59-2 produced β-carotene (all-E,9-Z,13-Z) and the cells having the pZea4 construct produced zeaxanthin. This confirms that we have cloned all the necessary genes of *Flavobacterium sp.* R1534 for the synthesis of zeaxanthin or their precursors (phytoene, lycopene and β-carotene). The production levels obtained are shown in table 1.

45

50

55

plasmid	h st	zeaxanthin	$\beta$ -carotene	lycopene
pLyco	<i>E. coli</i> JM109	ND	ND	0.05%
pBIKS(+-)clone59-2	"	ND	0.03%	ND
pZea4	"	0.033%	0.0009%	ND

15 Table 1: Carotenoid content of *E. coli* transformants, carrying the plasmids pLyco, pBIKS(+-)clone59-2 and pZea4, after 43 hours of culture in shake flasks. The values indicated show the carotenoid content in % of the total dry cell mass (200 ml). ND = not detectable.

25 **Examples 5**

**Carotenoid production in *B. subtilis***

In a first approach to produce carotenoids in *B. subtilis*, we cloned the carotenoid biosynthesis genes of 30 *Flavobacterium* into the Gram (+)/(-) shuttle vectors p602/22, a derivative of p602/20 [LeGrice, S.F.J., s.a.]. The assembling of the final construct p602-CARVEG-E, begins with a triple ligation of fragments Pvull-AvrII of pZea4(del654-3028) and the AvrII-EcoRI fragment from plasmid pBIKS(+-)clone6a, into the EcoRI and Scal sites of the vector p602/22. The plasmid pZea4(del654-3028) had been obtained by digesting pZea4 with SacI and EspI. The protruding and recessed ends were made blunt with Klenow enzyme and religated. Construct pZea4(del654-3028) lacks most of 35 the sequence upstream of crtE gene, which are not needed for the carotenoid biosynthesis. The plasmid p602-CAR has approx. 6.7 kb of genomic *Flavobacterium R1534* DNA containing besides all five carotenoid genes (approx. 4.9 kb), additional genomic DNA of 1.2 kb, located upstream of the crtZ translation start site and further 200 bp, located upstream of crtE transcription start. The crtZ, crtY, crtI and crtB genes were cloned downstream of the P<sub>N25/0</sub> promoter, a regulatable *E. coli* bacteriophage T5 promoter derivative, fused to a lac operator element, which is functional in *B. 40 subtilis* [LeGrice, S.F.J., s.a.]. It is obvious that in the p602CAR construct, the distance of over 1200 bp between the P<sub>N25/0</sub> promoter and the transcription start site of crtZ is not optimal and will be improved at a later stage. An outline of the p602CAR construction is shown in figure 17. To ensure transcription of the crtE gene in *B. subtilis*, the vegI promoter [Moran et al., Mol. Gen. Genet. 186, 339-346 (1982); LeGrice et al., Mol. Gen. Genet. 204, 229-236 (1986)] was introduced upstream of this gene, resulting in the plasmid construct p602-CARVEG-E. The vegI promoter, which originates from site of the veg promoter complex described by [LeGrice et al., s.a.] has been shown to be functional in *E. coli* [Moran et al., s.a.]. To obtain this new construct, the plasmid p602CAR was digested with Sall and HindIII, and the fragment containing the complete crtE gene and most of the crtB coding sequence, was subcloned into the Xhol and HindIII sites of plasmid p205. The resulting plasmid p205CAR contains the crtE gene just downstream of the Pvegl promoter. To reconstitute the carotenoid gene cluster of *Flavobacterium* sp. the following three pieces were isolated: 45 Pmel/HindIII fragment of p205CAR, the HinclI/XbaI fragment and the EcoRI/HindIII fragment of p602CAR and ligated into the EcoRI and XbaI sites of pBluescriptIIKS(+), resulting in the construct pBIKS(+-)-CARVEG-E. Isolation of the EcoRI-XbaI fragment of this latter plasmid and ligation into the EcoRI and XbaI sites of p602/22 gives a plasmid similar to p602CAR but having the crtE gene driven by the Pvegl promoter. All the construction steps to get the plasmid p602CARVEG-E are outlined in figure 18. *E. coli* TG1 cells transformed with this plasmid synthesized zeaxanthin. In 50 contrast *B. subtilis* strain 1012 transformed with the same constructs did not produce any carotenoids. Analysis of several zeaxanthin negative *B. subtilis* transformants always revealed, that the transformed plasmids had undergone severe deletions. This instability could be due to the large size of the constructs.

In order to obtain a stable construct in *B. subtilis*, the carotenoid genes were cloned into the Gram (+)/(-) shuttle vector pHP13 constructed by [Haima et al., s.a.]. The stability problems were thought to be omitted by 1) reducing the

size of the cloned insert which carries the carotenoid genes and 2) reversing the orientation of the crtE gene and thus only requiring one promoter for the expression of all five genes, instead of two, like in the previous constructs. Furthermore, the ability of cells transformed by such a plasmid carrying the synthetic *Flavobacterium* carotenoid operon (SFCO), to produce carotenoids, would answer the question if a modular approach is feasible. Figure 19 summarizes 5 all the construction steps and intermediate plasmids made to get the final construct pHp13-2PNZYIB-EINV. Briefly: To facilitate the following constructions, a vector pHp13-2 was made, by introducing a synthetic linker obtained with primer CS1 and CS2, between the HindIII and EcoRI sites of the shuttle vector pHp13. The intermediate construct pHp13-10 2CARVEG-E was constructed by subcloning the AflII-XbaI fragment of p602CARVEG-E into the AflII and XbaI sites of pHp13-2. The next step consisted in the inversion of crtE gene, by removing XbaI and AvrII fragment containing the original crtE gene and replacing it with the XbaI-AvrII fragment of plasmid pBIIKS(+)-PCRRBS crtE. The resulting plasmid was named pHp13-2CARZYIB-EINV and represented the first construction with a functional SFCO. The intermediate construct pBIIKS(+)-PCRRBS crtE mentioned above, was obtained by digesting the PCR product generated with primers #100 and #101 with Spel and SmaI and ligating into the Spel and SmaI sites of pBluescriptIIKS(+). In order to 15 get the crtZ transcription start close to the promoter P<sub>N25/0</sub> a triple ligation was done with the BamHI-Sall fragment of pHp13-2CARZYIB-EINV (contains four of the five carotenoid genes), the BamHI-EcoRI fragment of the same plasmid containing the P<sub>N25/0</sub> promoter and the EcoRI-Sall fragment of pBIIKS(+)-PCRRBS crtZ, having most of the crtZ gene preceded by a synthetic RBS. The aforementioned plasmid pBIISK(+)-PCRRBS crtZ was obtained by digesting the PCR product amplified with primers #104 and #105 with EcoRI and Sall and ligating into the EcoRI and Sall sites of pBluescriptIIKS(+). In the resulting vector pHp13-2PN25ZYIB-EINV, the SFCO is driven by the bacteriophage T5 promoter P<sub>N25/0</sub>, which should be constitutively expressed, due to the absence of a functional lac repressor in the construct 20 [Peschke and Beuk, J. Mol. Biol. 186, 547-555 (1985)]. *E. coli* TG1 cells transformed with this construct produced zeaxanthin. Nevertheless, when this plasmid was transformed into *B. subtilis*, no carotenoid production could be detected. Analysis of the plasmids of these transformants showed severe deletions, pointing towards instability problems, similar to the observations made with the aforementioned plasmids.

25

## Examples 6

### Chromosome Integration Constructs

30 Due to the instability observed with the previous constructs we decided to integrate the carotenoid biosynthesis genes of *Flavobacterium* sp. into the genome of *B. subtilis* using the integration/expression vector pXI12. This vector allows the constitutive expression of whole operons after integration into the levan-sucrase gene (sacB) of the *B. subtilis* genome. The constitutive expression is driven by the vegI promoter and results in medium level expression. The plasmid pXI12-ZYIB-EINV4 containing the synthetic *Flavobacterium* carotenoid operon (SFCO) was constructed as follows: the NdeI-HincII fragment of pBIISK(+)-PCRRBS crtZ was cloned into the NdeI and SmaI sites of pXI12 and the resulting plasmid was named pXI12-PCR crtZ. In the next step, the BstEII-Pmel fragment of pHp13-2PN25ZYIB-EINV was ligated to the BstEII-Pmel fragment of pXI12-PCR crtZ (see figure 20). *B. subtilis* transformed with the resulting construct pXI12-ZYIB-EINV4 can integrate the CAR genes either via a Campbell type reaction or via a reciprocal recombination. One transformant, BS1012::ZYIB-EINV4, having a reciprocal recombination of the carotenoid biosynthesis genes into the levan-sucrase gene was further analyzed (figure 21). Although this strain did not synthesize carotenoids, RNA analysis by Northern blots showed the presence of specific polycistronic mRNA's of 5.4 kb and 4.2 kb when hybridized to probe A (see figure 21, panel B). Whereas the larger mRNA has the expected message size, the origin of the shorter mRNA was unclear. Hybridization of the same Northern blot to probe B only detected the large mRNA fragment, pointing towards a premature termination of the transcription at the end of the crtB gene. The presence of a termination signal at this location would make sense, since in the original operon organisation in the *Flavobacterium* sp. R1534 genome, the crtE and the crtB genes are facing each other. With this constellation a transcription termination signal at the 5' end of crtB would make sense, in order to avoid the synthesis of anti-sense RNA which could interfere with the mRNA transcript of the crtE gene. Since this region has been changed considerably with respect to the wild type situation, the sequences constituting this terminator may also have been altered resulting in a "leaky" terminator. Western blot analysis using antisera against the different crt enzymes of the carotenoid pathway, pointed 40 towards the possibility that the ribosomal binding sites might be responsible for the lack of carotenoid synthesis. Out of the five genes introduced only the product of crtZ, the β-carotene hydroxylase was detectable. This is the only gene preceded by a RBS site, originating from the pXI12 vector, known to be functional in *B. subtilis*. Base pairing interactions between a mRNA's Shine-Dalgarno sequence [Shine and Delagarno, s. a.] and the 16S rRNA, which permits the ribosome to select the proper initiation site, have been proposed by [McLaughlin et al., J. Biol. Chem. 256, 11283-11291 45 (1981)] to be much more stable in Gram-positive organisms (*B. subtilis*) than in Gram-negative organisms (*E. coli*). In order to obtain highly stable complexes we exchanged the RBS sites of the Gram-negative *Flavobacterium* sp., preceding each of the genes crtY, crtI, crtB and crtE, with synthetic RBS's which were designed complementary to the 3' end 50 of the *B. subtilis* 16S rRNA (see table 2). This exchange should allow an effective translation initiation of the different 55 genes.

carotenoid genes in *B. subtilis*. The strategy chosen to construct this pXI12-ZYIB-EINV4MUTRBS2C, containing all four altered sites is summarized in figure 20. In order to facilitate the further cloning steps in pBluescriptIIKS(+), additional restriction sites were introduced using the linker obtained with primer MUT7 and MUT8, cloned between the Sall and HindIII sites of said vector. The new resulting construct pBIKS(+)-LINKER78 had the following restriction sites introduced: AvrII, PmlI, MspI, MnlI, BamHI and SphI. The general approach chosen to create the synthetic RBS's upstream of the different carotenoid genes, was done using a combination of PCR based mutagenesis, where the genes were reconstructed using defined primers carrying the modified RBS sites, or using synthetic linkers having such sequences. Reconstitution of the RBS preceding the crtI and crtB genes was done by amplifying the crtI gene with the primers MUT2 and MUT6, which include the appropriate altered RBS sites. The PCR-I fragment obtained was digested with 5 MnlI and BamHI and ligated into the MnlI and BamHI sites of pBIKS(+)-LINKER78. The resulting intermediate construct was named pBIKS(+)-LINKER78PCRI. Reconstitution of the RBS preceding the crtB gene was done using a small PCR fragment obtained with primer MUT3, carrying the altered RBS site upstream of crtB, and primer CAR17. The amplified PCR-F fragment was digested with BamHI and HindIII and sub cloned into the BamHI and HindIII sites of pBIKS(+)-LINKER78, resulting in the construct pBIKS(+)-LINKER78PCRF. The PCR-I fragment was cut out of pBIKS(+)-LINKER78PCRI with BamHI and SphI and ligated into the BamHI and SphI sites of pBIKS(+)-LINKER78PCRF. The resulting plasmid pBIKS(+)-LINKER78PCRF has the PCR-I fragment fused to the PCR-F fragment. This construct was cut with Sall and PmlI and a synthetic linker obtained by annealing of primer MUT9 and MUT10 was introduced. This latter step was done to facilitate the upcoming replacement of the original *Flavobacterium* RBS in the above mentioned construct. The resulting plasmid was named pBIKS(+)-LINKER78PCRFIA. Assembling of the synthetic 10 RBS's preceding the crtY and crtI genes was done by PCR, using primers MUT1 and MUT5. The amplified fragment PCR-G was made blunt end before cloning into the SmaI site of pUC18, resulting in construct pUC18-PCR-G. The next step was the cloning of the PCR-G fragment between the PCR-A and PCR-I fragments. For this purpose the PCR-G was isolated from pUC18-PCR-G by digesting with MnlI and PmlI and ligated into the MnlI and PmlI sites of pBIKS(+)-LINKER78PCRFIA. This construct contains all four fragments, PCR-F, PCR-I, PCR-G and PCR-A, assembled adjacent 15 to each other and containing three of the four artificial RBS sites (crtY, crtI and crtB). The exchange of the *Flavobacterium* RBS's preceding the genes crtY, crtI and crtB by synthetic ones, was done by replacing the HindIII-Sall fragment of plasmid pXI12-ZYIB-EINV4 with the HindIII-Sall fragment of plasmid pBIKS(+)-LINKER78PCRFIA. The resulting plasmid pXI12-ZYIB-EINV4 MUTRBS was subsequently transformed into *E. coli* TG1 cells and *B. subtilis* 1012. The production of zeaxanthin by these cells confirmed that the PCR amplified genes where functional. The *B. subtilis* strain obtained was named BS1012::SFCO1. The last *Flavobacterium* RBS to be exchanged was the one preceding the crtE gene. This was done using a linker obtained using primer MUT11 and MUT12. The wild type RBS was removed from pXI12-ZYIB-EINV4MUTRBS with NdeI and SphI and the above mentioned linker was inserted. In the 20 construct pXI12-ZYIB-EINV4MUTRBS2C all *Flavobacterium* RBS's have been replaced by synthetic RBS's of the consensus sequence AAAGGAGG- 7-8 N -ATG (see table 2). *E. coli* TG1 cells transformed with this construct showed that 25 also this last RBS replacement had not interferred

Table 2

40	<u>mRNA</u>	<u>nucleotide sequence</u>
	crtZ	<b>AAAGGAGGGUUUCAU<u>AUG</u>AGC</b>
45	crtY	<b>AAAGGAGGACACGUG<u>AUG</u>AGC</b>
	crtI	<b>AAAGGAGGCAU<u>UAG</u>AG<u>AUG</u>GU</b>
	crtB	<b>AAAGGAGGAU<u>CC</u>AU<u>CA</u><u>UAG</u>ACC</b>
50	crtE	<b>AAAGGAGGGUU<u>UCA</u><u>UAG</u>ACG</b>

55

<i>B. subtilis</i>	16S rRNA	3'- <b>UCUUUCCUCC</b> ACUAG
<i>E. coli</i>	16S rRNA	3'- AUUCCUCCACUAG

5

Table 2: Nucleotide sequences of the synthetic ribosome binding sites in the constructs pXI12-ZYIB-EINV4MUTRBS2C, pXI12-ZYIB-EINV4MUTRBS2CCAT and pXI12-ZYIB-EINV4 MUTRBS2CNEO. Nucleotides of the Shine-Dalgarno sequence preceding the individual carotenoid genes which are complementary to the 3' ends of the 16S rRNA of *B. subtilis* are shown in bold. The 3' ends of the 16S rRNA of *E. coli* is also shown as comparison. The underlined AUG is the translation start site of the mentioned gene.

with the ability to produce zeaxanthin. All the regions containing the newly introduced synthetic RBS's were confirmed by sequencing. *B. subtilis* cells were transformed with plasmid pXI12-ZYIB-EINV4MUTRBS2 and one transformant having integrated the SFCO by reciprocal recombination, into the levan-sucrase gene of the chromosome, was selected. This strain was named BS1012::SFCO2. Analysis of the carotenoid production of this strain show that the amounts zeaxanthin produced is approx. 40% of the zeaxanthin produced by *E. coli* cells transformed with the plasmid used to get the *B. subtilis* transformant. Similar was the observation when comparing the BS1012::SFCO1 strain with its *E. coli* counter part (30%). Although the *E. coli* cells have 18 times more carotenoid genes, the carotenoid production is only a factor of 2-3 times higher. More drastic was the difference observed in the carotenoid contents, between *E. coli* cells carrying the pZea4 construct in about 200 copies and the *E. coli* carrying the plasmid pXI12-ZYIB-EINV4MUTRBS2C in 18 copies. The first transformant produced 48x more zeaxanthin than the latter one. This difference seen can not only be attributed to the roughly 11 times more carotenoid biosynthesis genes present in these transformants. Contributing to this difference is probably also the suboptimal performance of the newly constructed SFCO, in which the overlapping genes of the wild type *Flavobacterium* operon were separated to introduce the synthetic RBS's. This could have resulted in a lower translation efficiency of the rebuild synthetic operon (e.g. due to elimination of putative translational coupling effects, present in the wild type operon).

In order to increase the carotenoid production, two new constructs were made, pXI12-ZYIB-EINV4MUTRBS2CNEO and pXI12-ZYIB-EINV4 MUTRBS2CCAT, which after the integration of the SFCO into the levan-sucrase site of the chromosome, generate strains with an amplifiable structure as described by [Janniere et al., Gene 40, 47-55 (1985)]. Plasmid pXI12-ZYIB-EINV4MUTRBS2CNEO has been deposited on May 25, 1995 at the DSM-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Germany) under accession No. DSM 10013. Such amplifiable structures, when linked to a resistance marker (e.g chloramphenicol, neomycin, tetracycline), can be amplified to 20-50 copies per chromosome. The amplifiable structure consist of the SFCO, the resistance gene and the pXI12 sequence, flanked by direct repeats of the sac-B 3' gene (see figure 22). New strains having elevated numbers of the SFCO could now be obtained by selecting for transformants with increased level of resistance to the antibiotic. To construct plasmid pXI12-ZYIB-EINV4MUTRBS2CNEO, the neomycin resistance gene was isolated from plasmid pBEST501 with PstI and SmaI and subcloned into the PstI and EcoO1091 sites of the pUC18 vector. The resulting construct was named pUC18-Neo. To get the final construct, the Pmel - AatII fragment of plasmid pXI12-ZYIB-EINV4MUTRBS2C was replaced with the SmaI-AatII fragment of pUC18-Neo, containing the neomycin resistance gene. Plasmid pXI12-ZYIB-EINV4MUTRBS2CCAT was obtained as follows: the chloramphenicol resistance gene of pC194 was isolated by PCR using the primer pair cat3 and cat4. The fragment was digested with EcoRI and AatII and subcloned into the EcoRI and AatII sites of pUC18. The resulting plasmid was named pUC18-CAT. The final vector was obtained by replacing the Pmel-AatII fragment of pXI12-ZYIB-EINV4MUTRBS2C with the EcoRI-AatII fragment of pUC18-CAT, carrying the chloramphenicol resistance gene. Figure 23 summarizes the different steps to obtain aforementioned constructs. Both plasmids were transformed into *B. subtilis* strain 1012, and transformants resulting from a

Campbell-type integration were selected. Two strains BS1012::SFCONEO1 and BS1012::SFCOCAT1 were chosen for further amplification. Individual colonies of both strains were independently amplified by growing them in different concentrations of antibiotics as described in the methods section. For the cat gene carrying strain, the chloramphenicol concentrations were 60, 80, 120 and 150 µg/ml. For the neo gene carrying strain, the neomycin concentrations were 5 160 and 180 µg/ml. In both strains only strains with minor amplifications of the SFCO's were obtained. In daughter strains generated from strain BS1012::SFCONEO1, the resistance to higher neomycin concentrations correlated with the increase in the number of SFCO's in the chromosome and with higher levels of carotenoids produced by these cells. A different result was obtained with daughter strains obtained from strain BS1012::SFCOCAT1. In these strains an increase up to 150 µg chloramphenicol/ml resulted, as expected, in a higher number of SFCO copies in the chromosome. 10

### Example 7

#### Construction of CrtW containing plasmids and use for carotenoid production

15 *Polymerase chain reaction based gene synthesis.* The nucleotide sequence of the artificial crtW gene, encoding the β-carotene β-4-oxygenase of *Alcaligenes* strain PC-1, was obtained by back translating the amino acid sequence outlined in [Misawa, 1995], using the BackTranslate program of the GCG Wisconsin Sequence Analysis Package, Version 8.0 (Genetics Computer Group, Madison, WI, USA) and a codon frequency reference table of *E. coli* (supplied by the Bach Translate Program). The synthetic gene consisting of 726 nucleotides was constructed basically according to the method described by [Ye, 1992]. The sequence of the 12 oligonucleotides (crtW1 - crtW12) required for the synthesis are shown in Figure 25. Briefly, the long oligonucleotides were designed to have short overlaps of 15-20 bases, serving as primers for the extension of the oligonucleotides. After four cycles a few copies of the full length gene should be present which is then amplified by the two terminal oligonucleotides crtW15 and crtW26. The sequences for these two short oligonucleotides are for the forward primer crtW15 (5'-TATATCTAGAcatatgTCCGGTCTGAAA CCGG-3') and for the reverse primer crtW26 (5'-TATAAgattccacgtTCA AGCACGACCACCGGTTTACG-3'), where the sequences matching the DNA templates are underlined. Small cap letters show the introduced restriction sites (*Nde*I for the forward primer and *Eco*RI and *Pml*I for the reverse primer) for the latter cloning into the pALTER-Ex2 expression vector.

20 *Polymerase chain reaction.* All twelve long oligonucleotides (crtW1-crtW12; 7 nM each) and both terminal primers (crtW15 and crtW26; 0.1 mM each) were mixed and added to a PCR reaction mix containing Expand™ High Fidelity polymerase (Boehringer, Mannheim) (3.5 units) and dNTP's (100 mM each). The PCR reaction was run for 30 cycles with the following profile: 94 °C for 1 min, 50 °C for 2 min and 72 °C for 3 min. The PCR reaction was separated on a 1% agarose gel, and the band of approx. 700 bp was excised and purified using the glass beads method (Geneclean Kit, Bio101, Vista, CA, USA). The fragment was subsequently cloned into the *Sma*I site of plasmid pUC18, using the 25 Sure-Clone Kit (Pharmacia, Uppsala, Sweden). The sequence of the resulting crtW synthetic gene was verified by sequencing with the Sequenase Kit Version 1.0 (United States Biochemical, Cleveland, OH, USA). The crtW gene constructed by this method was found to contain minor errors, which were subsequently corrected by site-directed mutagenesis.

30 *Construction of plasmids.* Plasmid pBIIKS(+)·CARVEG-E (see also Example 5) (Figure 26) contains the carotenoid biosynthesis genes (crtE, crtB, crtY, crtI and crtZ) of the Gram (-) bacterium *Flavobacterium* sp. strain R1534 WT (ATCC 21588) [Pasamontes, 1995 #732] cloned into a modified pBluescript II KS(+) vector (Stratagene, La Jolla, USA) carrying site I of the *B. subtilis* veg promoter [LeGrice, 1986 #806]. This constitutive promoter has been shown to be functional in *E. coli*. Transformants of *E. coli* strain TG1 carrying plasmid pBIIKS(+)·CARVEG-E synthesise zeaxanthin. Plasmid pALTER-Ex2-crtW was constructed by cloning the *Nde*I - *Eco*RI restricted fragment of the synthetic crtW gene 35 into the corresponding sites of plasmid pALTER-Ex2 (Promega, Madison, WI). Plasmid pALTER-Ex2 is a low copy plasmid with the p15a origin of replication, which allows it to be maintained with *Cole*I vectors in the same host. Plasmid pBIIKS-crtEBIYZW (Figure 26) was obtained by cloning the *Hind*III-*Pml*I fragment of pALTER-Ex2-crtW into the *Hind*III and the blunt end made *Mlu*I site obtained by a fill in reaction with Klenow enzyme, as described elsewhere in [Sambrook, 1989 #505]. Inactivation of the crtZ gene was done by deleting a 285 bp *Nsi*I-*Nsi*I fragment, followed by a fill in 40 reaction and religation, resulting in plasmid pBIIKS-crtEBIY[ΔZ]W. Plasmid pBIIKS-crtEBIY[ΔZ]W carrying the non-functional genes crtW and crtZ, was constructed by digesting the plasmid pBIIKS-crtEBIY[ΔZ]W with *Nde*I and *Hpa*I, and subsequent self religation of the plasmid after filling in the sites with Klenow enzyme. *E. coli* transformed with this 45 plasmid had a yellow-orange colour due to the accumulation of β-carotene. Plasmid pBIIKS-crtEBIY[ΔW] has a truncated crtW gene obtained by deleting the *Nde*I - *Hpa*I fragment in plasmid pBIIKS-crtEBIYZW as outlined above. Plasmids pALTER-Ex2-crtEBIY[ΔZ]W and pALTER-Ex2-crtEBIY[ΔW], were obtained by isolating the *Bam*HI-*Xba*I fragment from pBIIKS-crtEBIY[ΔZ]W and pBIIKS-crtEBIY[ΔW], respectively and cloning them into the *Bam*HI and *Xba*I sites of pALTER-Ex2. The plasmid pBIIKS-crtW was constructed by digesting pBIIKS-crtEBIYZW with *Nsi*I and *Sac*I, and self-relinquishing the plasmid after recessing the DNA overhangs with Klenow enzyme. Figure 27 compiles the 50 relevant inserts of all the plasmids used in this paper.

5 *Carotenoid analysis.* *E. coli* TG-1 transformants carrying the different plasmid constructs were grown for 20 hours in Luria-Broth medium supplemented with antibiotics (ampicillin 100 µg/ml, tetracyclin 12.5 µg/ml) in shake flasks at 37°C and 220 rpm. Carotenoids were extracted from the cells with acetone. The acetone was removed in vacuo and the residue was re dissolved in toluene. The coloured solutions were subjected to high-performance liquid chromatography (HPLC) analysis which was performed on a Hewlett-Packard series 1050 instrument. The carotenoids were separated on a silica column Nucleosil Si - 100, 200 x 4 mm, 3m. The solvent system included two solvents: hexane (A) and hexane/THF, 1:1 (B). A linear gradient was applied running from 13 to 50 % (B) within 15 minutes. The flow rate was 1.5 ml/min. Peaks were detected at 450 nm by a photo diode array detector. The individual carotenoid pigments were identified by their absorption spectra and typical retention times as compared to reference samples of chemically pure carotenoids, prepared by chemical synthesis and characterised by NMR, MS and UV-Spectra. HPLC analysis of the pigments isolated from *E. coli* cells transformed with plasmid pBIKS-crtEBIYZW, carrying besides the carotenoid biosynthesis genes of *Flavobacterium* sp. strain R1534, also the crtW gene encoding the β-carotene ketolase of *Alcaligenes* PC-1 [Misawa, 1995 #670] gave the following major peaks identified as: β-cryptoxanthin, astaxanthin, adonixanthin and zeaxanthin, based on the retention times and on the comparison of the absorbance spectra to given reference samples of chemically pure carotenoids. The relative amount (area percent) of the accumulated pigment in the *E. coli* transformant carrying pBIKS-crtEBIYZW is shown in Table 3 ["CRX": cryptoxanthin; "ASX": astaxanthin; "ADX": adonixanthin; "ZCN": zeaxanthin; "ECM": echinenone; "MECH": 3-hydroxyechinenone, "CXN": canthaxanthin]. The Σ of the peak areas of all identified carotenoids was defined as 100%. Numbers shown in Table 3 represent the average value of four independent cultures for each transformant. In contrast to the aforementioned results, *E. coli* transformants carrying the same genes but on two plasmids namely, pBIKS-crtEBIYZ[ΔW] and pALTER-Ex2-crtW, showed a drastic drop in adonixanthin and a complete lack of astaxanthin pigments (Table 3), whereas the relative amount of zeaxanthin (%) had increased. Echinenone, hydroxyechinenone and canthaxanthin levels remained unchanged compared to the transformants carrying all the crt genes on the same plasmid (pBIKS-crtEBIYZΔW). Plasmid pBIKS-crtEBIYZ[ΔW] is a high copy plasmid carrying the functional genes of crtE, crtB, crtY, crtI, crtZ of *Flavobacterium* sp. strain R1534 and a truncated, non-functional version of the crtW gene, whereas the functional copy of the crtW gene is located on the low copy plasmid pALTER-Ex2-crtW. To analyze the effect of overexpression of the crtW gene with respect to the crtZ gene, *E. coli* cells were co-transformed with plasmid pBIKS-crtW carrying the crtW gene on the high copy plasmid pBIKS-crtW and the low copy construct pALTER-Ex2-crtEBIYZ[ΔW], encoding the *Flavobacterium* crt genes. Pigment analysis of these transformants by HPLC monitored the presence of β-carotene, cryptoxanthin, astaxanthin, adonixanthin, zeaxanthin, 3-hydroxyechinenone and minute traces of echinenone and canthaxanthin (Table 3). Transformants harbouring the crtW gene on the low copy plasmid pALTER-Ex2-crtW and the genes crtE, crtB, crtY and crtI on the high copy plasmid pBIKS-crtEBIY[ΔZW] expressed only minor amounts of canthaxanthin (6%) but high levels of echinenone (94%), whereas cells carrying the crtW gene on the high copy plasmid pBIKS-crtW and the other crt genes on the low copy construct pALTER-Ex2-crtEBIY[ΔZW], had 78.6 % and 21.4 % of echinenone and canthaxanthin, respectively (Table 3).

Table 3

plasmids	CRX	ASX	ADX	ZCN	ECH	HECH	CXN
pBIKS-crtEBIYZW	1.1	2.0	44.2	52.4	< 1	< 1	< 1
pBIKS-crtEBIYZ[ΔW] + pALTER-Ex2-crtW	2.2	-	25.4	72.4	< 1	< 1	< 1
pBIKS-crtEBIY[ΔZ]W	-	-	-	-	66.5	-	33.5
pBIKS-crtEBIY[ΔZW] + pBIKS-crtW	-	-	-	-	94	-	6

50 **Claims**

1. A DNA sequence comprising one or more DNA sequences selected from the group consisting of:
  - a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous;
  - b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) or a DNA sequence which is substantially homologous;

c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI) or a DNA sequence which is substantially homologous;

5 d) a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which is substantially homologous;

e) a DNA sequence which encodes the  $\beta$ -carotene hydroxylase of *Flavobacterium* sp. R1534 (crtZ) or a DNA sequence which is substantially homologous.

10 2. A DNA sequence as claimed in claim 1 comprising the following DNA sequences:

a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous, and

15 b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) or a DNA sequence which is substantially homologous, and

c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI) or a DNA sequence which is substantially homologous.

20 3. A DNA sequence as claimed in claim 1 comprising the following DNA sequences:

a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous, and

25 b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) or a DNA sequence which is substantially homologous, and

c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI) or a DNA sequence which is substantially homologous, and

30 d) a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which is substantially homologous.

35 4. A DNA sequence as claimed in claim 1 comprising the following DNA sequences:

a) a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE) or a DNA sequence which is substantially homologous, and

40 b) a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB) or a DNA sequence which is substantially homologous, and

c) a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI) or a DNA sequence which is substantially homologous, and

45 d) a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which is substantially homologous, and

e) a DNA sequence which encodes the  $\beta$ -carotene hydroxylase of *Flavobacterium* sp. R1534 (crtZ) or a DNA sequence which is substantially homologous.

50 5. A DNA sequence as claimed in claim 4 which comprises in addition to the DNA sequences specified in claim 4 a DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of *Alcaligenes* strain PC-1 (crtW) or a DNA sequence which is substantially homologous.

55 6. A DNA sequence as claimed in claim 3 which comprises in addition to the DNA sequences specified in claim 3 a DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of *Alcaligenes* strain PC-1 (crtW) or a DNA sequence which is substantially homologous.

7. A vector comprising the DNA sequence of claim 1.
8. A vector comprising the DNA sequence of claim 2.
- 5 9. A vector comprising the DNA sequence of claim 3.
- 10 10. A vector comprising the DNA sequence of claim 4.
11. A vector comprising the DNA sequence of claim 5.
12. A vector comprising the DNA sequence of claim 6
13. A cell which is transformed by the DNA sequence of claim 1 or the vector of claim 7.
- 15 14. A cell which is transformed by the DNA sequence of claim 2 or the vector of claim 8.
- 15 15. A cell which is transformed by the DNA sequence of claim 3 or the vector of claim 9.
16. A cell which is transformed by the DNA sequence of claim 4 or the vector of claim 10.
- 20 17. A cell which is transformed by the DNA sequence of claim 5 or the vector of claim 11.
18. A cell which is transformed by the DNA sequence of claim 6 or the vector of claim 12.
- 25 19. A cell which is transformed by the DNA sequence of claim 4 or the vector of claim 10 and a second DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of Alcaligenes strain PC-1 (crt W) or a DNA sequence which is substantially homologous or a second vector which comprises a DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of Alcaligenes strain PC-1 (crt W) or a DNA sequence which is substantially homologous.
- 30 20. A cell which is transformed by the DNA sequence of claim 3 or the vector of claim 9 and a second DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of Alcaligenes strain PC-1 (crt W) or a DNA sequence which is substantially homologous or a second vector which comprises a DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of Alcaligenes strain PC-1 (crt W) or a DNA sequence which is substantially homologous.
- 35 21. The cell of any one of claims 13 to 20 which is a prokaryotic cell.
22. The cell of claim 21 which is E. coli.
23. The cell of claim 21 which is a Bacillus strain.
- 40 24. The cell of any one of claims 13 to 20 which is an eukaryotic cell.
25. The cell of claim 24 which is a yeast cell.
- 45 26. The cell of claim 24 which is a fungal cell.
27. A process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing a cell as claimed in any one of claims 13 to 26 under suitable culture conditions and isolating the desired carotenoid or a mixture of carotenoids from such cells or the culture medium and, in case only one carotenoid is desired separating it by methods known in the art from other carotenoids which might also be present.
- 50 28. A process as claimed in claim 27 for the preparation of lycopene by culturing a cell as claimed in claim 14.
29. A process as claimed in claim 27 for the preparation of  $\beta$ -carotene by culturing a cell as claimed in claim 15.
- 55 30. A process as claimed in claim 27 for the preparation of echinenone by culturing cells as claimed in claim 18 or 20.
31. A process as claimed in claim 27 for the preparation of canthaxanthin by culturing cells as claimed in claim 18.

**EP 0 747 483 A2**

32. A process as claimed in claim 27 for the preparation of zeaxanthin by culturing cells as claimed in claim 17 or 19.
33. A process as claimed in claim 27 for the preparation of adonixanthin by culturing cells as claimed in claim 17 or 19.
- 5 34. A process as claimed in claim 27 for the preparation of astaxanthin by culturing cells as claimed in claim 17.
35. A process for the preparation of a food or feed composition characterized therein that after a process as claimed in any one of claims 27 to 34 has been effected the carotenoid or carotenoid mixture is added to food or feed.

10

15

20

25

30

35

40

45

50

55

Fig. 1

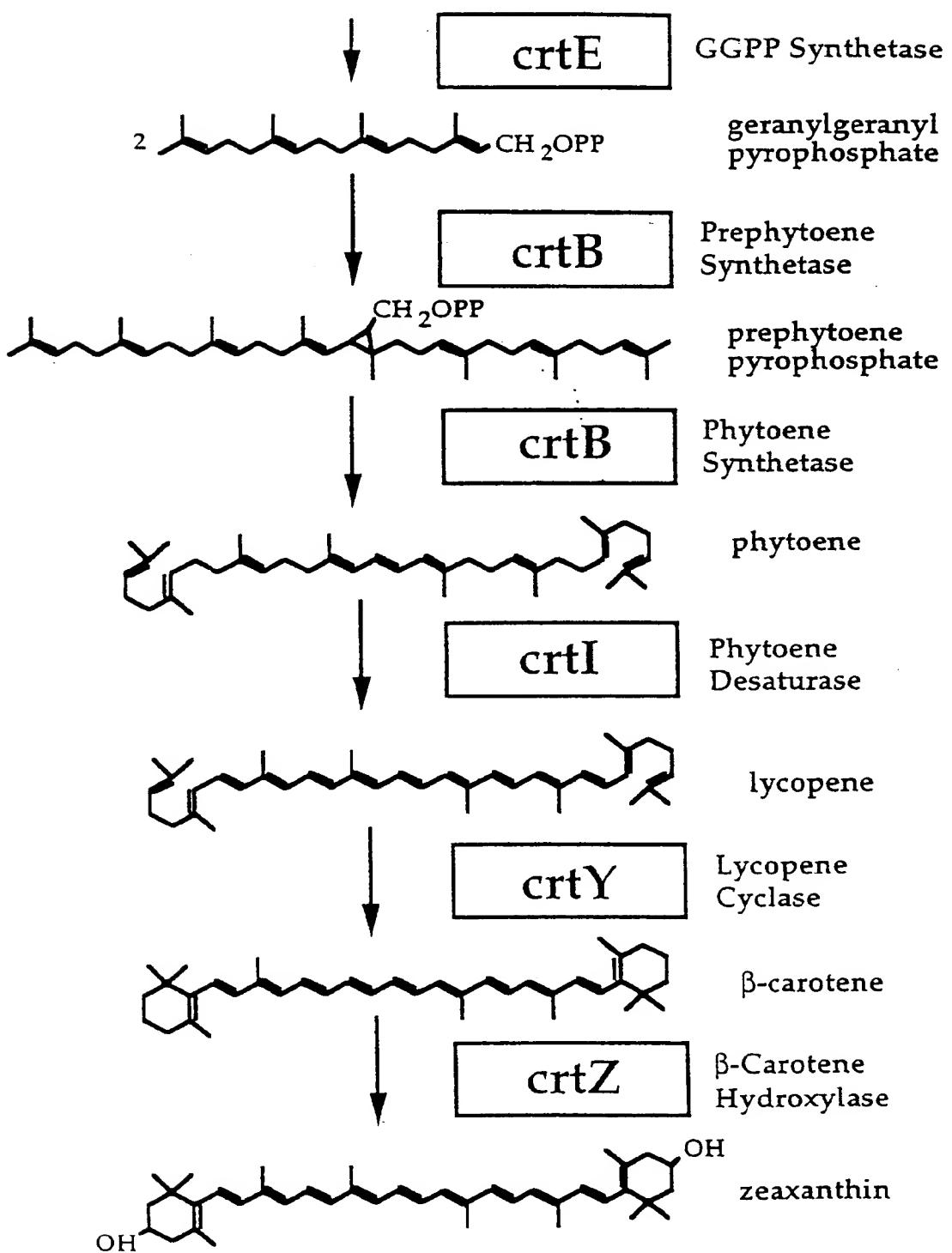
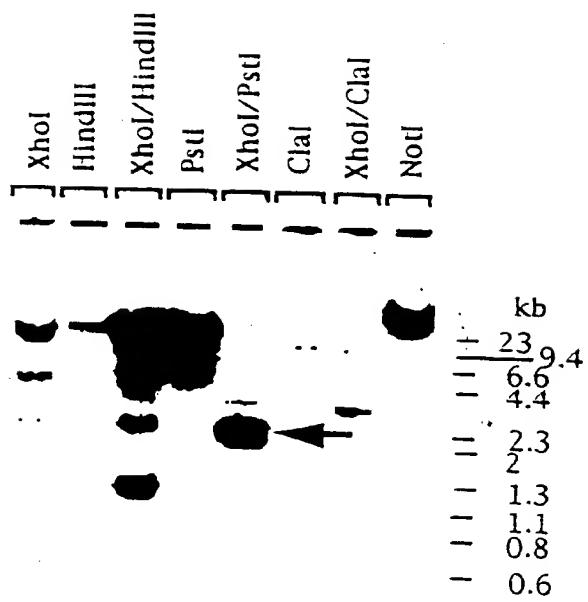


Fig. 2



**Fig. 3**

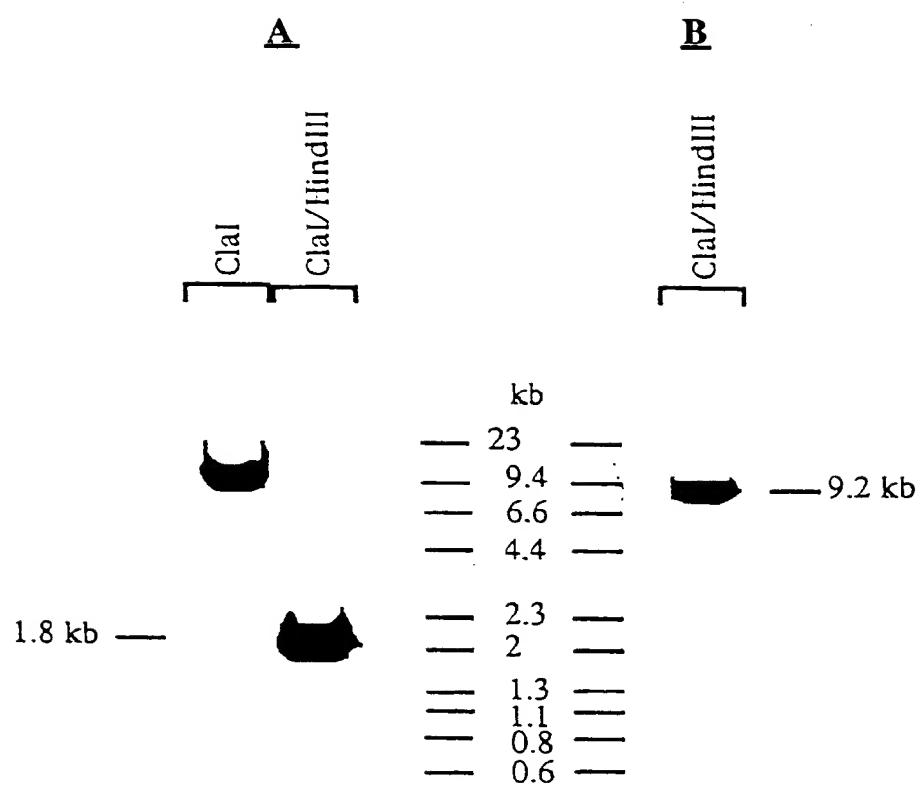


Fig. 4

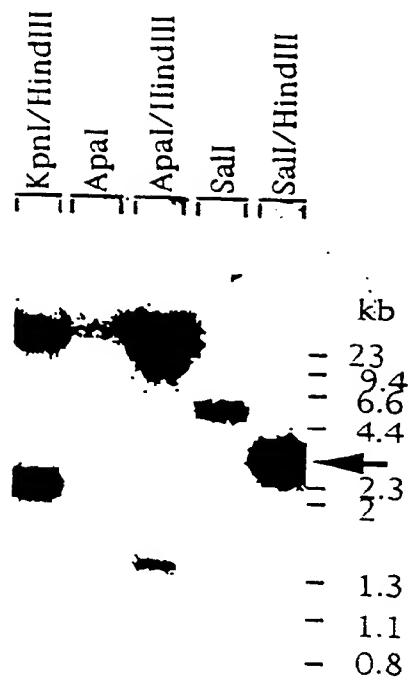


Fig. 5

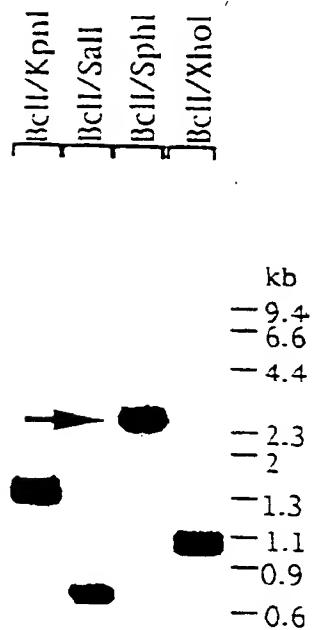


Fig. 6

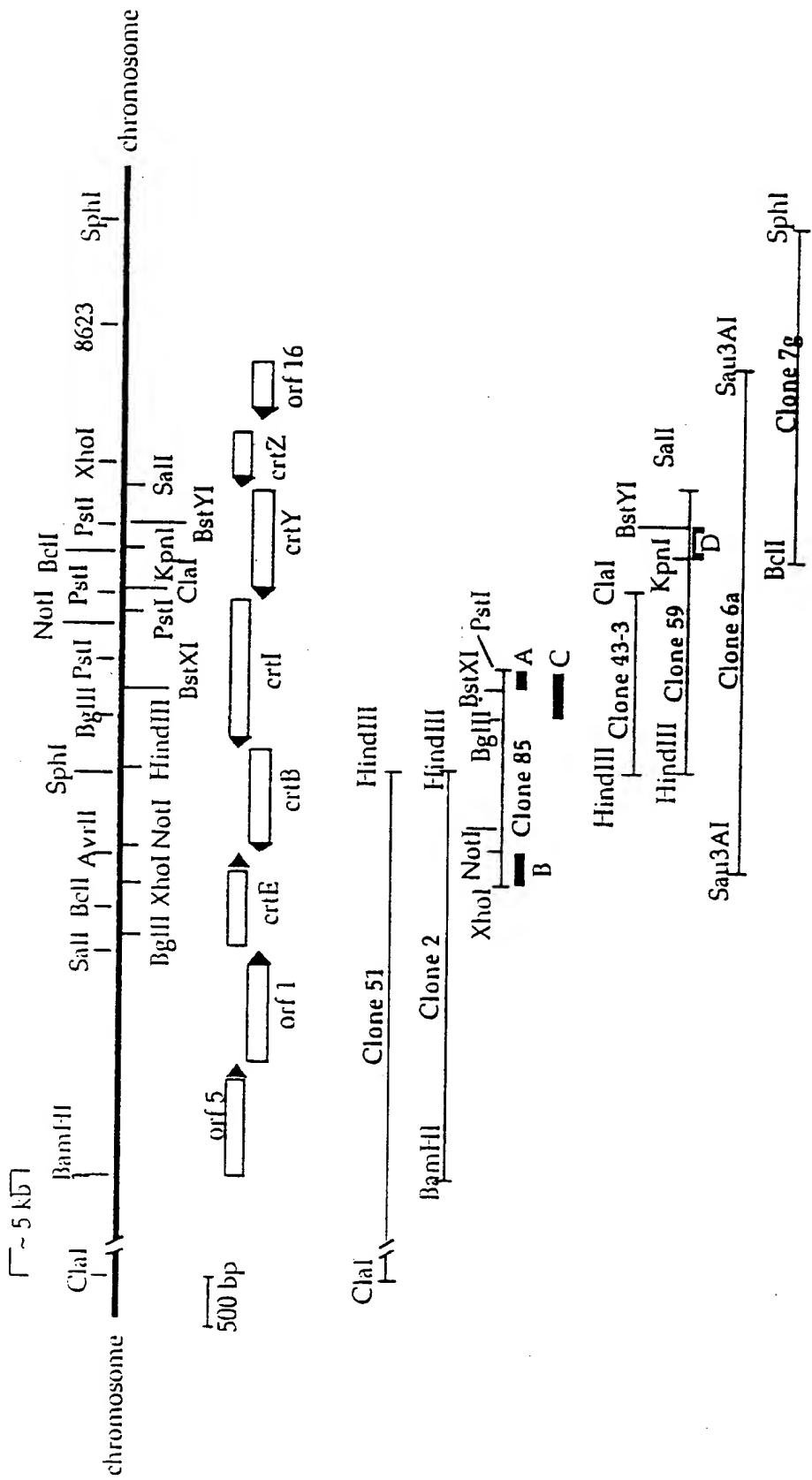


Fig. 7/1

GGATCCGGCCCTGGCGATGATCGAGCGCCGCGCTTGGGATCGTC  
 1 -----+-----+-----+-----+-----+-----+-----+-----+  
 CCTAGGGCGAACCGGCGAACGGCGAACGGCGAACGGCGAACGGCG  
 51 -----+-----+-----+-----+-----+-----+-----+-----+  
 orf-5 -> D P R L A V R D Q Q P P L R I G Q  
 AGCATATCCCCATGAAACCGCAAGCCACGAGCCAGCGCCAGATC  
 51 -----+-----+-----+-----+-----+-----+-----+-----+  
 TCGTAGGGGTACTCGGCGCCGTGCGTGGCGCGGGGTCTAG  
 100 -----+-----+-----+-----+-----+-----+-----+-----+  
 H H P H E P Q R T T Q R A P Q I  
 GGGGGCTTCAACGCGGCTGGCCATCATGGGAGGCGCCGGCG  
 101 -----+-----+-----+-----+-----+-----+-----+-----+  
 CCCGGCGACGGCTGGCGCTACCGGGTAAAGGGCTTCAGGGCGCG  
 G R V Q H G M R H H R E G P R R B  
 TGGGGCGCTGGCCATTCGAGAAACTCGCAAGCTCGCTGGCGCG  
 151 -----+-----+-----+-----+-----+-----+-----+-----+  
 ACCCGCGCAACGGGTAGGGCTTGTGCGTGGCAAGCGCGTCC  
 G A R A H S E E L A A C P L R K V  
 TCGGGCGCTGGCCGCTATGCCATTCGGATCGGCGCCATGCCGCT  
 201 -----+-----+-----+-----+-----+-----+-----+-----+  
 AGCGCGCTTCAGCGGCGCTATAGCTTCTCGCTAGCGCG  
 A P D R A V F R C S D G P D A R  
 GGGGGCGCCGGCGGCGGCGGCGGCGGCG  
 251 -----+-----+-----+-----+-----+-----+-----+-----+  
 CGGGGGCGGCGGCGGCGGCGGCGGCG  
 G P A L P R R H Q R I A H E P F R  
 AGATGATGCTGCTGATCAATGGCCGTCATGCGAACGATCAACGGATCC  
 301 -----+-----+-----+-----+-----+-----+-----+-----+  
 TCTACTACAGGACTAAGTACGGTACGGTAACTGTTTGGCTAGGCTAGG  
 350 -----+-----+-----+-----+-----+-----+-----+-----+  
 D D V L I H G P S L Q N R S P I L  
 TGTGGCGCTGATGCGCTTGGCGATGGCGCGGGCTTGGATGGCGCGA  
 351 -----+-----+-----+-----+-----+-----+-----+-----+  
 ACAGGGGACTAACGAACTCGGCGCTCCGCGATCTAACGGCG  
 400 -----+-----+-----+-----+-----+-----+-----+-----+  
 S R D G I V C N A P R A R M A R  
 AGGATCAGGGGGAGAGACATGGAAATCGAGGACGGCTTGTGCT  
 401 -----+-----+-----+-----+-----+-----+-----+-----+  
 TCCATAGTCGCCCGCTCTGTGACCTTACCTCCGAAAGACCGA  
 R I K Q Q R D M R I R G R V F V V  
 CACGGGGCGACATCGGCGCTGGCGGCGCTGGGGGGATGCTGGCG  
 451 -----+-----+-----+-----+-----+-----+-----+-----+  
 GTGGCGCGGGCTAGGGCGAGGCGCGGCGCTAGGGCGCG  
 T G A A S G Q L G A A A S A R M L A Q  
 AAGGGGGCGGCGCTGGCGCGCTGGCGCGCTGGCGCG  
 501 -----+-----+-----+-----+-----+-----+-----+-----+  
 TCCGGGGCGGGCTGGCGCGCTGGCGCGCTGGCGCG  
 Q G A K V V I A D L A E P K D A  
 CGGAGGGCGGGCTGGCGCGCTGGCGCGCTGGCGCG  
 551 -----+-----+-----+-----+-----+-----+-----+-----+  
 GGGCTGGCGCGCTGGCGCGCTGGCGCGCTGGCGCG  
 P E G A V B A A C D V T D A T A

Fig. 7/2

Fig. 7/3

1201	G A N R T P M G A F Q G D L A A N	GGATGCCCGAACCTTGCGCGGAGCGGATCCGGCCGCTGAGGGCGT CGCGCTTACCGTGGCTGGCTAACCCGATTCAGGAGGATCTGGCGCAT	1250	V V A G G M E S M S N A P Y L L P	CAAGGGCGCCCTGGGATGGCAATGGCGATGGCGATGGCGATCAAC CAGAGCGGGCGCCCTAACCTCTGACCGCATCTGAAACGCCCTAACCTGGCG	1550
1251	D A P T L G A D A I R A A L N G L	TGTCGGCGGACATGGTGGAGGTGGTGTGGCTGGCTGGCTGGCG ACAGCCGGCTGTCACCTGGCTCCAGACTGGCAAGGAGGGCGCG	1300	K A R S G M R M G H D R V L D H K	TGTCGGCGGCTGGGCTGGAGGAGCGGCTATGGCAAGGGCCCTGATGGCG ACAGAGCTGGCCGACACTCTCGGATATGGTTCGGCGCTAACCCGATCTACCGCG	1551
1301	S P D M V D E V L W M G C V L A A	TGGTCTCGAGGGTGGAGGAGCGGCTATGGCAAGGGCCCTGATGGCG ACAGAGCTGGCCGACACTCTCGGATATGGTTCGGCGCTAACCCGATCTACCGCG	1350	F L D G L E D A Y D K G R L M G	TGGTCTCGAGGGTGGAGGAGCGGCTATGGCAAGGGCCCTGATGGCG ACAGAGCTGGCCGACACTCTCGGATATGGTTCGGCGCTAACCCGATCTACCGCG	1601
1351	G Q Q A P A R Q A A L G A G L P	GGCCAGGGTCAGGCAACCGGACACGGTCAAGGGCGCCCTGGCCGACTGCC CGGGTCCAGTGGCTGGCGCTGGCACTCTGGCGCGGAAACCGGCGCTGAGCG	1400	T F A E D C A G D H G F T R E A Q	GGCCAGGGTCAGGCAACCGGACACGGTCAAGGGCGCCCTGGCCGACTGCC CGGGTCCAGTGGCTGGCGCTGGCACTCTGGCGCGGAAACCGGCGCTGAGCG	1651
1401	L 9 T G T T I N E N C C G M K	CCTGCGGACGGGACGACCACTGAGGAGATGGCGATGGCGATGGCGCATGA CGACAGCTGGCCCTGCTGGTGGTAGTGGCTCATACGGCGTAGCGGGCGTAG	1450	D D Y A L T S L A R A Q D A I A S	GGCCAGGGTCAGGCAACCGGACACGGTCAAGGGCGCCCTGGCCGACTGCC CGGGTCCAGTGGCTGGCGCTGGCACTCTGGCGCGGAAACCGGCGCTGAGCG	1701
1451	A G G C G G C A T G C T G C G C C A T G C C C T G A T C G C C G C C G G A T C G C G C C G C G G A T C G C G C A C G C G A A G T C C G G G C C T A C G A C C G G T A C G A C C T A G G C G C C C T A G C G G G C G T A G	1500		GGCCAGGGTCAGGCAACCGGACACGGTCAAGGGCGCCCTGGCCGACTGCC CGGGTCCAGTGGCTGGCGCTGGCACTCTGGCGCGGAAACCGGCGCTGAGCG	1751	

Fig. 7/4

Fig. 7/5

2401	CGT GACCCGATGGCAACCGATGGGCGATGCCGATGCA	2450	CGT GACCCGATGGCAACCGATGGGCGATGCCGATGCA	2500	CTGACCGCAACGAGCTGGCAACCGATGGGCGATGCCG	2550	GACTGGCTGATCCGGGATCTGGCTATCGGGGATCTGG	2600	GATCTGGGATCTGGCTGACCTGGGGATCTGGGCTG	2650	CTGACCGCAACGAGCTGGCAACCGATGCCGATGCA	2700	AACGTTTCGCCCCCTGCTGAGCTGGCGGAACTGGGG	2750	CTGCGATCTGGCGGCTGAGCTGGCGGAACTGGGG
2701	CGCTTGGCGCTGGCAACCGATGCCGATGCA	2751	CGCTTGGCGCTGGCAACCGATGCCGATGCA	2801	CTGGCGGCTGGCGCTGGCGCTGGCGCTGGCGCTGG	2850	CTGGCGGCTGGCGCTGGCGCTGGCGCTGGCGCTGG	2901	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	2950	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3000	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG		
2801	CGCTTGGCGCTGGCAACCGATGCCGATGCA	2851	CGCTTGGCGCTGGCAACCGATGCCGATGCA	2901	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	2950	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3000	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3050	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3100			
2850	CGCTTGGCGCTGGCAACCGATGCCGATGCA	2900	CGCTTGGCGCTGGCAACCGATGCCGATGCA	2950	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3000	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3050	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3100	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3150			
2900	CGCTTGGCGCTGGCAACCGATGCCGATGCA	2950	CGCTTGGCGCTGGCAACCGATGCCGATGCA	3000	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3050	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3100	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3150	GGCGCGGCGCTGGCGCTGGCGCTGGCGCTGGCG	3200			

Fig. 7/6

3001	A P K D A A G I E R E Q D L K T G	CGCCCGCCCGGGGATCGAACTGGGCTCGAGATCTGCGATTTAGGCTCTGG GCGAGACAGGGCGCCGGCTAGCTAGACAGCTTAATTCGAGAACCC	3050	V L F V A G L E E N L S I I K G L D	ACAGGGCGAGCGAGGAGCTCCATGGCTCGAGCTGGCTGGCTGG TGTTCCGCTCTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	3100	K A E T E Q L M A F G R Q L G R	GCTTCCAGCTCTATGAGCACTGGCTGCACTGCGATAGAGCCAG CAGAAGGTGGGATGATCTGGCTGGCTGGCTGGCTGGCTGGCTGG	3150	V F Q S Y D D L L D V I G D K A S	CAACGGCAAGGATACTGGCGCGGCGGCGGCGGCGGCG GTGCGCGCGCTCCATGGCGCGCTGGCTGGCGCTGGCG	3200	T G K D T A R D T A A P G P K G G	GCCTGATGGGGTGGGAAGTGGCGGCGACGCTGGCGCGCTGGCG CGGAGCTACCGCTGGCGCTGGCGCTGGCGCTGGCG	3250	I K A A K S T I R S I R Q R Y A E P	CAACCTGGGATGGCGCGCTGGCGCTGGCGCTGGCGCTGGCG GTCGGACCGCTGGCGCTGGCGCTGGCGCTGGCG	3300	3350	3400	3450	3500	3550	3600
3101	A Q L D E L M R T R L F R G G	GGAGATCGGGGACCTGGCCGGCTGCTGGCGATGAACTGGCG CGCTAGCGCTGGCGACAGGGCGCTGGCGCTGGCGCTGGCG	3151	Q I A D L L A R V L P H D I R R S	GGGGCTTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG GGGGCTTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG	3201	A *	A A R P R T W L G D R S I E G	3251	G R L R S A A A D L G G R A L L G	GGCTCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG CTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG	3301	3351	3400	3450	3500	3550	3600						
3201	A Q L D E L M R T R L F R G G	GGAGATCGGGGACCTGGCCGGCTGCTGGCGATGAACTGGCG CGCTAGCGCTGGCGACAGGGCGCTGGCGCTGGCGCTGGCG	3251	G R L R S A A A D L G G R A L L G	GGGGCTTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG GGGGCTTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG	3301	3351	3400	3450	3500	3550	3600												
3301	A Q L D E L M R T R L F R G G	GGAGATCGGGGACCTGGCCGGCTGCTGGCGATGAACTGGCG CGCTAGCGCTGGCGACAGGGCGCTGGCGCTGGCGCTGGCG	3351	G R L R S A A A D L G G R A L L G	GGGGCTTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG GGGGCTTGGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG	3401	3451	3500	3550	3600														

Fig. 7/7

Fig. 7/8

4201    GCGATAGGCGTACGCGTATCTCGGGATTCGGGCGGGCGGCGATCGCTGGCG  
 CGTATCGTACTGGCATAGGCGCTACGGCGCTAGGCGTACTGCGACCG  
 A Y L N V T D E R I G P P M L K A

4250    GCGCTGGGAGCTTTCGGAACCTTGCGCTGGCGATGGGACTGCG  
 GCGGAACCGCTTCGAAACCTGGGACTGGCTAACGGGAGGCGCTTCAGGC  
 A Q A F S Q S G Q A I A A E S T

4300    CGCTCAGATCGCTCATCGAACCGCCAGTCGCACTGACCTGGCG  
 GCGACTCTAACGAGTACCGCTGGGCTAACGGCTTCAGGCGCTAACGGC  
 A T L D T M  
 <-- artB

4350    CGCTCAGATCGCTCATCGAACCGCCAGTCGCACTGACCTGGCG  
 GCGACTCTAACGAGTACCGCTGGGCTAACGGCTTCAGGCGCTAACGGC  
 T A K A S G V V G P I G A G P H T

4400    CGCGCCCGCTGGCGCTGCACAGCAACCGCGATGGCGCGACCGGATGGCG  
 ACCGAAACCGCGACGCGCTGGCGCTAACGGCTAACGGCTAACGGCG  
 G A G V I Y R N P I A R D R N H P

4450    CGCGACCGCGCTGGCGCTAACGGCTAACGGCTAACGGCG  
 CGCTTGGTGGCGCTAACGGCTAACGGCTAACGGCTAACGGCG  
 R F W A S Q T L I P P E V S F A S

4501    CGTATGGCGGAGCTGGCGCTAACGGCTAACGGCG  
 GCGCTGGCGCTAACGGCTAACGGCTAACGGCG  
 G R H H A S L E T S F D A P P S F I R

4551    CTGACGGCTAGGTGCTTGCCTAGGTGGGATGGCGGGCTCCAGTTC  
 GACTGCGATGTCACAGGCTAACGGCTAACGGCG  
 S V T L H K R L D P I A R R E E L E

4600    CTGCGAGATCGCTGGCGCTAACGGCTAACGGCTAACGGCG  
 GAGCTTTCAGGAGCGCTAACGGCTAACGGCG  
 E F I R E A Y G P A K A E W D V

4650    CGCGCCGGCGCTGGCGCTAACGGCTAACGGCG  
 GCGCGCCGGCTAACGGCTAACGGCTAACGGCG  
 D A R G L H P V P A L V Y H T S W

4700    CGCGCCGGCGCTGGCGCTAACGGCTAACGGCG  
 GCGAGCGCCGGCTAACGGCTAACGGCG  
 G E P A L S P D T V C P S H L Y M

4751    CGAGAAATGGCGCGCTAACGGCTAACGGCG  
 GCTCTTGGCGCTAACGGCTAACGGCG  
 S F D D P L I R P G N F I E N V L

Fig. 7/9

4801	CCCTGTAGCCGCCGAAATGACGCTGTGGGGCCACCTGGG GGAAACATCGCCCGCCCTTACTGACACCAACCGGTCCAAAGCCCC	4850	G K Y R P G F I V S H H A L N E P	4851	CGCTTGGACGGCCGAATGAGCAAGGACATGACCTGGCTG GGAAACATCGCCCGCCCTTACTGACACCAACCGGTCCAAAGCCCC	4900	R F P L A V M G A V L Q N T G G K	4901	CGCTTGGACGGCCGAATGAGCAAGGACATGACCTGGCTG GGAAACATCGCCCGCCCTTACTGACACCAACCGGTCCAAAGCCCC
4950	CGGGTTCAAGATCGCCCGCTTGGTGCCTGGGGGATGGCA GGCCGAAAGTCCTAGGCCCGAAACCGGGGACCCATACCGGTCT	4950	R N L I A A K T R G R R T H G L	4951	GGTGGGATAGCTGCTGCTACGTCGCTTGGGGGGGGGGGG CCACCGCTATGAGCTAGGAGGGGGGGGGGGGGGGGGGGGGGG	5000	38	5000	GGTGGGATAGCTGCTGCTACGTCGCTTGGGGGGGGGGGGGG CCACCGCTATGAGCTAGGAGGGGGGGGGGGGGGGGGGGGGGGGG
5050	CGGAAACTGGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGG GGCTTGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5050	R L Q R G D L L T V G T A R D G E	5051	CGGTGGGATGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGG CCACGGCTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5100	LD R Y S H M V D G N S A V T D A	5051	CGGTGGGATGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGG CCACGGCTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
5101	GCTGAGACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG CGAGCTTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5101	5150	5151	GCTGAGACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG CGAGCTTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5200	5200	5201	GAGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GCTGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
5250	CGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GCTGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5250	S S T S F P N G Q V L L T H F S	5251	AGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG TCGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5300	5300	5301	AGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG TCGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
5350	TCGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GCTGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5350	F A Q R L H P D Q I P R A V N S H	5351	V B R Y A Q L R M L A P A N L M	5400	5400	5401	CTGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GACGG
5450	CTGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GACGG	5450	T D I R T V R A N L L I T G G L	5451	Q C L K L F P T F G L K L Y G E	5500	5500	5501	CTGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG GACGG

Fig. 7/10

Fig. 7/11

Fig. 7/12

6601 6650 6651 6701 6751 6801 6851

ATCGTCAAGGCTCGGAATGCTTACCGCGATGCAACCGCTG  
TAGAGGTCCAGCGCTGACCCATAGGTGGCTCTACTGTGGCACT  
D D L T A V H T N W R L D V G Q  
CGACCCGATCAGCGCCCGCTCGATGAGCCATAGCTGCTGCTAGG  
CGTCCGGGCTGTCGGGCGGCGGCGGCGGCGGCGGCG  
L L G I L A G A K I S G Y G T T L  
CGCGCGGAAATGGTCCGCAACCGACCGACCGTCTGATTCGGCTCATTCGGCG  
GGCGCGCTAACCGCCCTTGGCTGAGGACTAGCGCTAGGCGCTAAGGGCGC  
R R S H D P F A V E Q D T W E G R  
ACGATGGCGACAGGGCGCCATTGGCGGAAAGATCGCTCT  
TGCTTAACCGCTGTCGGGCGCTGGCTTGGCGACAGCA  
R I P S L R A L W E P S I D T D  
CGCGAGGACCGGTGCTGGCTGGAGGGCGGACCGCGGCGCTCGAGCTC  
CGCGCTGTCGACAGACAGACGGCTCGCCGGCTGGCGCGCTGG  
R C S W T H Q D S P G S R A D L M  
ACATAGCGCCATCGCGCTGGCTGGCGACCGCGATCGCGC  
TGCTTAACCGCTAGGCGGAGACCGCAACCGCTGGCTTGGCG  
V I R A D P R R D R V A L A I L A

6900 6950 6951 7001 7051 7101 7151

ACCGGAGCGCCCGGCGGCGGCGGCGGCGGCGGCG  
TGGCTGGCTGTCGACCGCTGGCTGGCTGGCTGGCTGGCTGG  
G S L G A G A I L L D H S M  
ATCCGCGCCCTTCGGGGTCTCGCCACCGCGCCGGCG  
TGGCGCGCCAGACCGAGAGACGCTGGCTGGCGCGCG  
D A G E R D K L L A G S R K L E  
CGCTGAGCGCTGTCGACCGCGCCCGCGCGCGCG  
GAGACTCGCGACAGCGCTGGCGCGCTGGCTGGCTGGCG  
A K L S D V S P A W I F G F S V C  
TTCCTGCGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCG  
AAGAGCGCGCGCTGGCGCGCGCTGGCTGGCTGGCTGGCG  
N E R G H V A H H M R H A Q Y V R  
ACGAGATAGCGCGCGCTGGCGCGCGCGCGCGCG  
TGCTTCATGGCGCGCGCGCGCGCGCGCGCGCGCGCG  
R L Y Q R K P V Y R F P W R Q H  
CGAGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCG  
GGTTCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCG  
V L Q D H L F Y Y I L Q Y C T V G

6950 7000 7050 7100 7150

ATCCGCGCCCTTCGGGGTCTCGCCACCGCGCCGGCG  
TGGCGCGCCAGACCGAGAGACGCTGGCTGGCGCGCG  
D A G E R D K L L A G S R K L E  
CGCTGAGCGCTGTCGACCGCGCCCGCGCGCGCG  
GAGACTCGCGACAGCGCTGGCGCGCTGGCTGGCG  
A K L S D V S P A W I F G F S V C  
TTCCTGCGCGCTGGCTGGCTGGCTGGCTGGCTGGCG  
AAGAGCGCGCGCTGGCGCGCGCTGGCTGGCTGGCG  
N E R G H V A H H M R H A Q Y V R  
ACGAGATAGCGCGCGCTGGCGCGCGCGCGCGCG  
TGCTTCATGGCGCGCGCGCGCGCGCGCGCGCGCG  
R L Y Q R K P V Y R F P W R Q H  
CGAGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCG  
GGTTCGCGCTGGCGCGCGCGCGCGCGCGCGCGCG  
V L Q D H L F Y Y I L Q Y C T V G

Fig. 7/13

Fig. 7/14

7801	GGCGGATGCCCAATGCCCTGAGCCGCTCAGGCCAGGGGATGCCCTC GGGCTACGGGTTAGCCGAGCTGGCCAGAGTCCGCTAGGGAG	7850	A R E E P S D A V E V R F G L T E	8101	GGGGCTCTTCGGGCTGCCCCACCTGCCGAAACCGAGGGTTTC CGGGGAGAAGGGGAGGGGAGGGCTGGAGGGCTTGGCTGAGAG	8150
7851	TGGGGGGGAAATTCGAGGAAACGGGCTGGGGATCGGGCA AGGGGGGGCTAAAGCTCTGGCTTGCGCAAGGCCCTAGGGCT	7900	A G T D V V L S G P A C E V A A	8151	CGGGGGGGGAAATTCGAGGAAACGGGCTGGGGATCGGGCA GGGGGGGGCTAAAGCTCTGGCTTGCGCAAGGCCCTAGGGCT	8200
7901	G G G A I E L V F L R D P D P D G	7950	A A A P M L V . A L L A A A K S P W	8201	CGGGGGGGGGAAATTCGAGGAAACGGGCTGGGGATCGGGCA GGGGGGGGCTAAAGCTCTGGCTTGCGCAAGGCCCTAGGGCT	8250
7951	V A A G P I P T E D L P R A N R H	8000	A A A P M L V . A L L A A A K S P W	8251	ATGGGGGAAATTCGAGGAAACGGGCTGGGGATCGGGCA TACACGGCTACTGGGGCTAAAGCTGGCTTGCGCAAGGCCCTA	8300
8001	I H R I V G T E D A F V M D L P I	8050	N P L I P S S P A S I R S V R R M	8301	CCTGGTCCGCTCTTCAGGGGAAATTCGAGGAAACGGGCTGGGGAT GGGGGGGGCTAAAGCTCTGGCTTGCGCAAGGCCCTAGGGCT	8350
8051	L T N R M W R S V P Q P S E Y I	8100	R T G T M <-- orf-16	8351	ACGGGATTCGGCATCGAGAAGGAGCCGAGCTGGGGAAATCGGGCTGG TGTGCTTAAACCCGCTGGCTCTTCCTGAGCTGGCTTGAGCTACT	8400

Fig. 7/15

8401 CACCAAGTCGAGAGCCGGAAATGCGAGCACCTCGATATGGATGACCA  
 +-----+-----+-----+-----+-----+-----+-----+  
 GTGCTCGAGCTCTTGGCCCTACTGCTCGTGGACCTATACCTACTCTG 8450

8451 CGTCCCTCGGGCTGGCGAAATGTTGGCGAAACCGGGAAAAGGCCCTTGGC  
 +-----+-----+-----+-----+-----+-----+-----+  
 GCAAGGACCCCGCCCGCTTCACACCGCTTTCACCGGAAACCG 8500

8501 CTTGTGAAACACTTGAACGGGGGGACGCCACGGCAnnCGTCCAGATG  
 +-----+-----+-----+-----+-----+-----+-----+  
 GAAACGGCTTGGAAACTGGAACCGCCTGGCTGGTGGTnGCGTAC 8550

8551 CTCGATAACTCTGGCATTCAGATCGGGAAATGGGGGTGnGnGCGCTT  
 +-----+-----+-----+-----+-----+-----+-----+  
 GAGCTACTGAGGCCCTAGCTAGGCTAGCCCTnCCCCCAACnGnGACGGAAA 8600

8601 CnnnCGGTTGATCGACGGACCTC  
 +-----+-----+-----+-----+-----+-----+  
 GnnnCCAAAGCTAGCTGCTGGAG 8625

Fig. 8

1 MTPKQQFPLR DLVEIRLAQI SGQFGVVSAP LGAAMSDAAL SPGKRFRAVL  
51 MLMVAESSGG VCDAMVDAAC AVEMVHAASL IFDDMPCMDD ARTRRGQPAT  
101 HVAHGEGRAV LAGIALITEA MRILGEARGA TPDQRARLVA SMSRAMGPVG  
151 LCAGQDLDLH APKDAAGIER EQDLKTGVLF VAGLEMLSII KGLDKAETEQ  
201 LMAFGRQLGR VFQSYDDLLD VIGDKASTGK DTARDTAAPG PKGGLMAVGQ  
251 MGDVAQHYRA SRAQLDELMR TRLFRGGQIA DLLARVLPHD IRRSA

Fig.9

1 MTDLTATSEA AIAQGSQSFA QAAKLMPPGI REDTVMLYAW CRHADDVIDG  
51 QVMGSAPEAG GDPQARLGAL RADTLAALHE DGPMSPFCAA LRQVARRHDF  
101 PDLWPMDLIE GFAMDVADRE YRSLDDVLEY SYHVAGVVGV MMARVMGVQD  
151 DAVLDRACDL GLAFQLTNIA RDVIDDAAIG RCYLPADWLA EAGATVEGPV  
201 PSDALYSVII RLLDAAEPYY ASARQGLPHL PPRCAWSIAA ALRIYRAIGT  
251 RIRQGGPEAY RQRISTSKAA KIGLLARGGL DAAASRLRGG EISRDGLWTR  
301 PRA

Fig. 10

1 MSSAIVIGAG FGGLALAIRL QSAGIATTIV EARDKPGGRA YVWNDQGHVF  
51 DAGPTVVTDP DSLRELWALS GQPMERDVTL LPVSPFYRLT WADGRSFYEV  
101 NDDDELIROV ASFNPADVDG YRRFHDXAEE VYREGYLKLG TPPFLKLGQM  
151 LNAAPALMRL QAYRSVHSMV ARFIQDPHLR QAFSFHTLLV GGNPFSTSSI  
201 YALIHALERR GGVWFAKGGT NQLVAGMVAL FERLGGTLLL NARVTRIDTE  
251 GDRATGVTLL DGRQLRADTV ASNGDVMHSY RDLLGHTRRG RTKAAILNRQ  
301 RWSMSLFVLH FGLSKRPENL AHHSVIFGPR YKGLVNEIFN GPRLPDDFSM  
351 YLHSPCVTDP SLAPEGMSTH YVLAPVPHLG RADVDWEAEA PGYAERIFEE  
401 LERRAIPDLR KHLTVSRIFS PADFSTELSA HHGSAFSVEP ILTQSAWFRP  
451 HNRDRAIPNF YIVGAGTHPG AGIPGVVGSA KATAQVMLSD LAVA

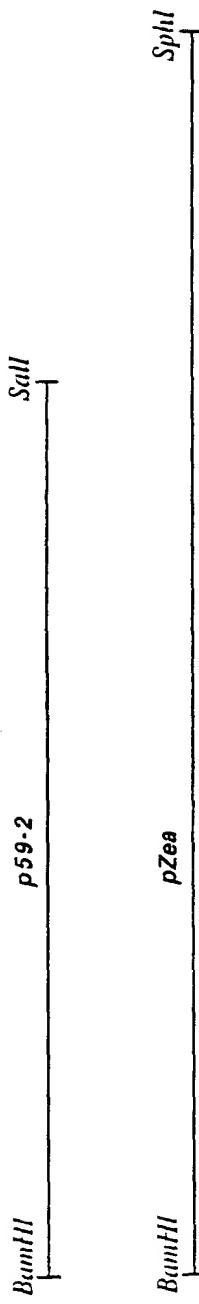
Fig.11

1 MSHDLLIAGA GLSGALIALA VRDRRPDARI VMLDARSGPS DQHTWSCHDT  
51 DLSPEWLARL SPIRRGEWTD QEVAFPDHSR RLTTGYGSIE AGALIGLLQG  
101 VDLRWNTHVA TLDDTGATLT DGSRIEAACV IDARGAVETP HLTVGFQKFW  
151 GVEIETDAPH GVERPMIMDA TVPQMDGYRF IYLLPFSPTR ILIEDTRYSD  
201 GGDLDGALA QASLDYAARR GWTGQEMRRE RGILPIALAH DAIGFWRDHA  
251 QGAVPVGLGA GLFHPVTGYS LPYAAQVADA IAARDLTTAS ARRAVRGWAI  
301 DRADRDRFLR LLNRMLFRGC PPDRRYRLLQ RFYRLPQPLI ERFYAGRLTL  
351 ADRLRIVTGR PPIPLSQAVR CLPERPLLQE RA

Fig. 12

1 MSTWAAILTV ILTVAAMELT AYSVHRWIMH GPLGWGWHKS HHDEDHDHAL  
51 EKNDLYGVIF AVISIVLFAI GAMGSDLAWW LAVGVTCYGL IYYFLHDGLV  
101 HGRWPFRYVP KRGYLRRVYQ AHRMHHAVHG RENCVSFGFI WAPSVDSDLKA  
151 ELKRSGALLK DREGADRNT

Fig. 13



construct	<i>crtE</i>	<i>crtB</i>	<i>crtI</i>	<i>crtY</i>	<i>crtZ</i>	<i>carotene</i>
<i>pLycO</i>	+	+	+	-	-	LYCOPENE
<i>p59-2</i>	+	+	+	+	-	$\beta$ -CARTOTENE
<i>pZea 4</i>	+	+	+	+	+	ZEAXANTHIN

Fig. 14

#100: 5' tataatcatattaaaggacaattacatATGACGCCAAGCAGCAGCAATT 3'  
 Spel RBS NdeI → crtE

#101: 5' TATATACCCGGGTCAGCCGCGACGGCCTGTGG 3'  
 SmaI

#104: 5' tataatcatattaaaggaggaattacatATGAGCACITGGGCCGCAATCC 3'  
 EcoRI RBS NdeI → crtZ

#105: 5' GTTTCAGCTCTGCCTTGAGGC 3'

MUT1: 5' GCGAAGGGGCGGATCGCAATACttaaggaaggacacttatATGAGCCATGATCTGCTGATCG 3'  
 PmlI → crtZ → crtY

MUT2: 5' GCCCCCTGCTGCAGGAGAGAGCtTaaggaggcattagATGAGTTCCGCCATCGTCATCG 3'  
 MunI → crtY → crtI

MUT3: 5' GGTCATGCTGTCGGACCTGGCCGTCGtTaaggaggaacattatATGACCGATCTGACGGCAGTTCC 3'  
 BamHI → crtI → crtB

MUT5: 5' ATATATatttccttccttttcaaaGCTCTCCTGCAGCAGGG 3'  
 MunI ← crtY

MUT6: 5' atgattttccttccttttcaaaGCGACGGCCAGGTCCGACAGC 3'  
 BamHI ← crtI

CAR17 5' CAGAACCCATCACCTGCCCGTC 3'

ca3: 5' CGCGAATTCTGGCCGGCAATAGTTACC 3'  
 EcoRI

ca4: 5' GTCACATGCATGCATTTACGGCTCAAAGCATGTGACCCTTTCAAATAACGGGGCAGG 3'  
 SphI SacI AatII

१५

CS1:	5'	<u>AGCTTGGATCCTTAAGTACTCTAGAGTTAACG</u>	EcoRI
CS2:	3'	ACCTAGGAATTCATGAGATCTCAAATTGCTTAA	3'
		<u>Scal</u>	
		<u>XbaI</u>	
		<u>HindIII</u>	

**MUT7:** 5' TCGACCTAGGCACCGTGACCGTCAATTGGATCGCATGCAAGCTT 3'  
**MUT8:** 3' GGATCCGGACTGGCAGTTAACCTAGGGTACGGTACGTTGAACTAG 5'

SalI \_\_\_\_\_  
 3'-  
 MUTY: 5' gtgtcccttttacGATATTGGATCCGGCCCTTCGGGTCCCTAGCTCTGGCTG  
 MUT10: 3' cacaggaaaatgtgATAAACGCTAGGGGGAAAGGCCAGGAACCTCCGACAGCT  
 5'-  
 RBS 

Fig. 16

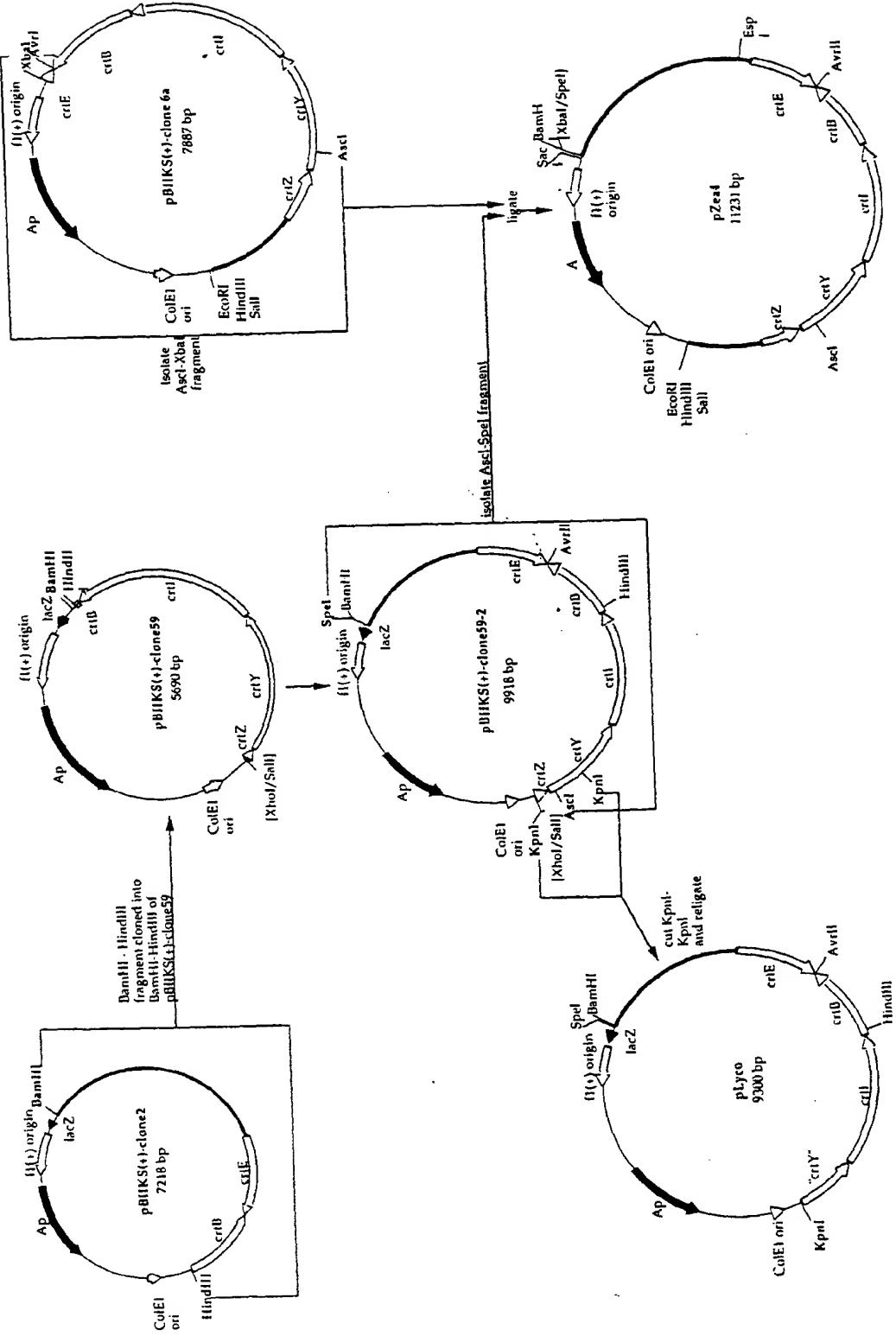


Fig. 17

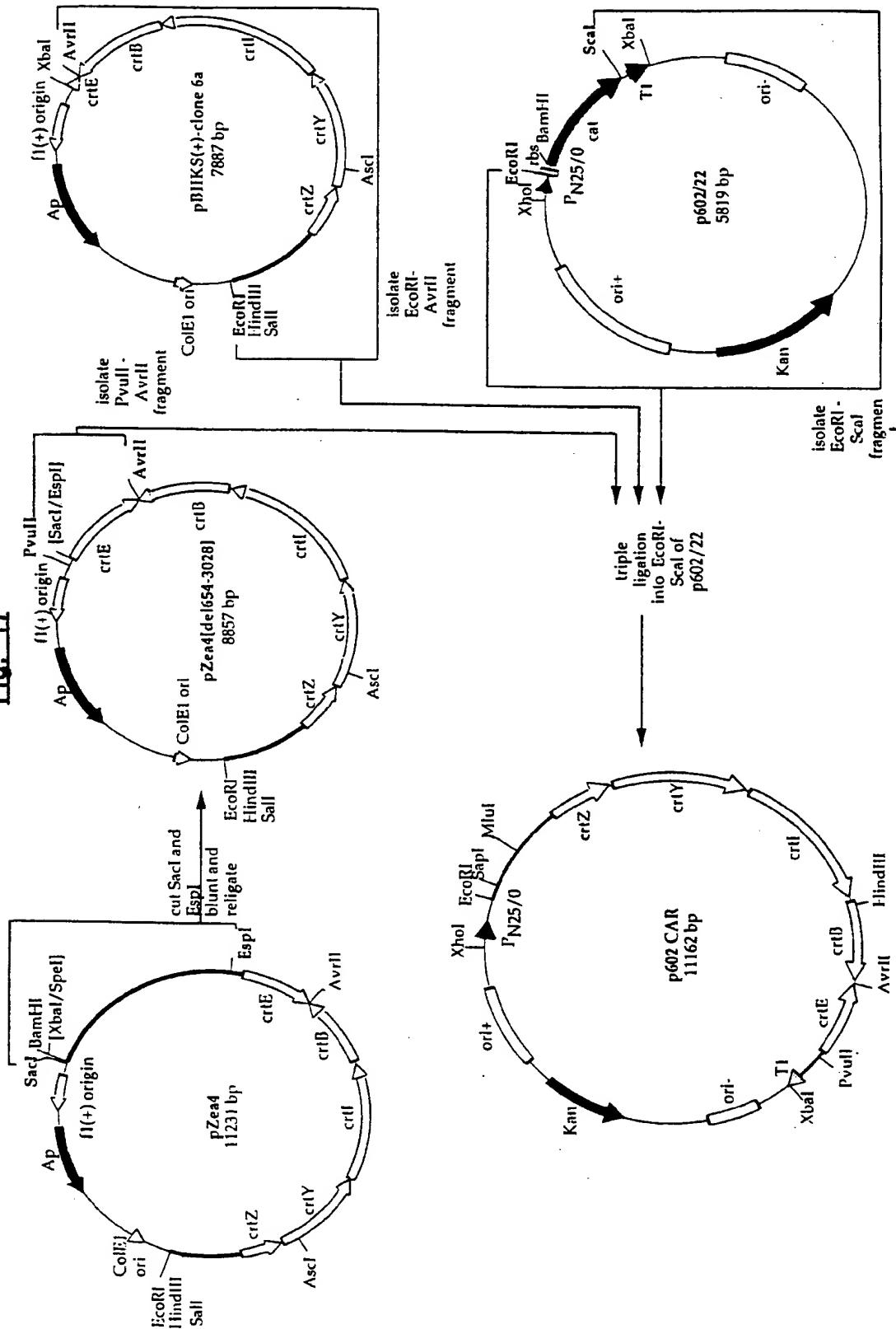


Fig. 18

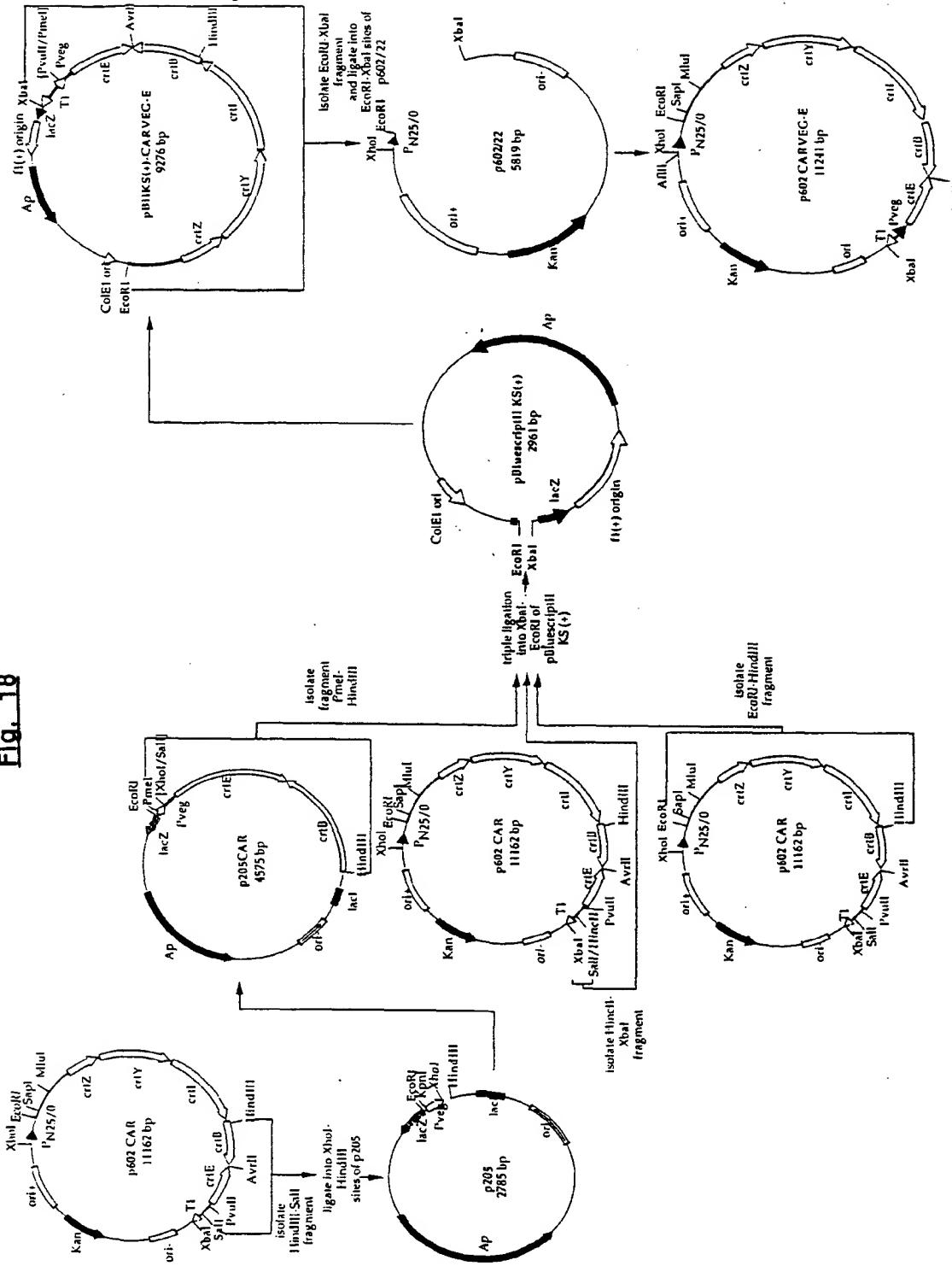


Fig. 19

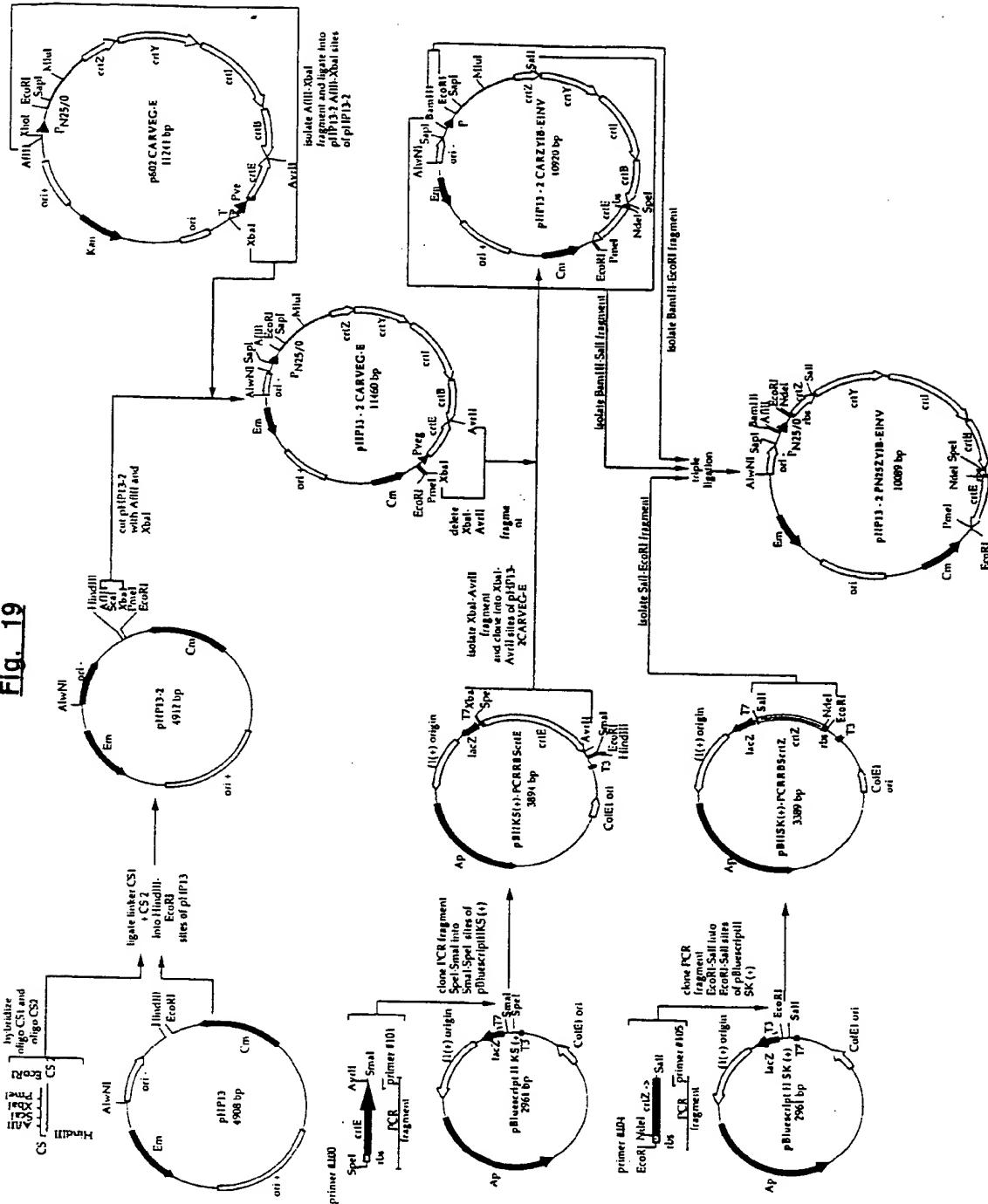


Fig. 20/1

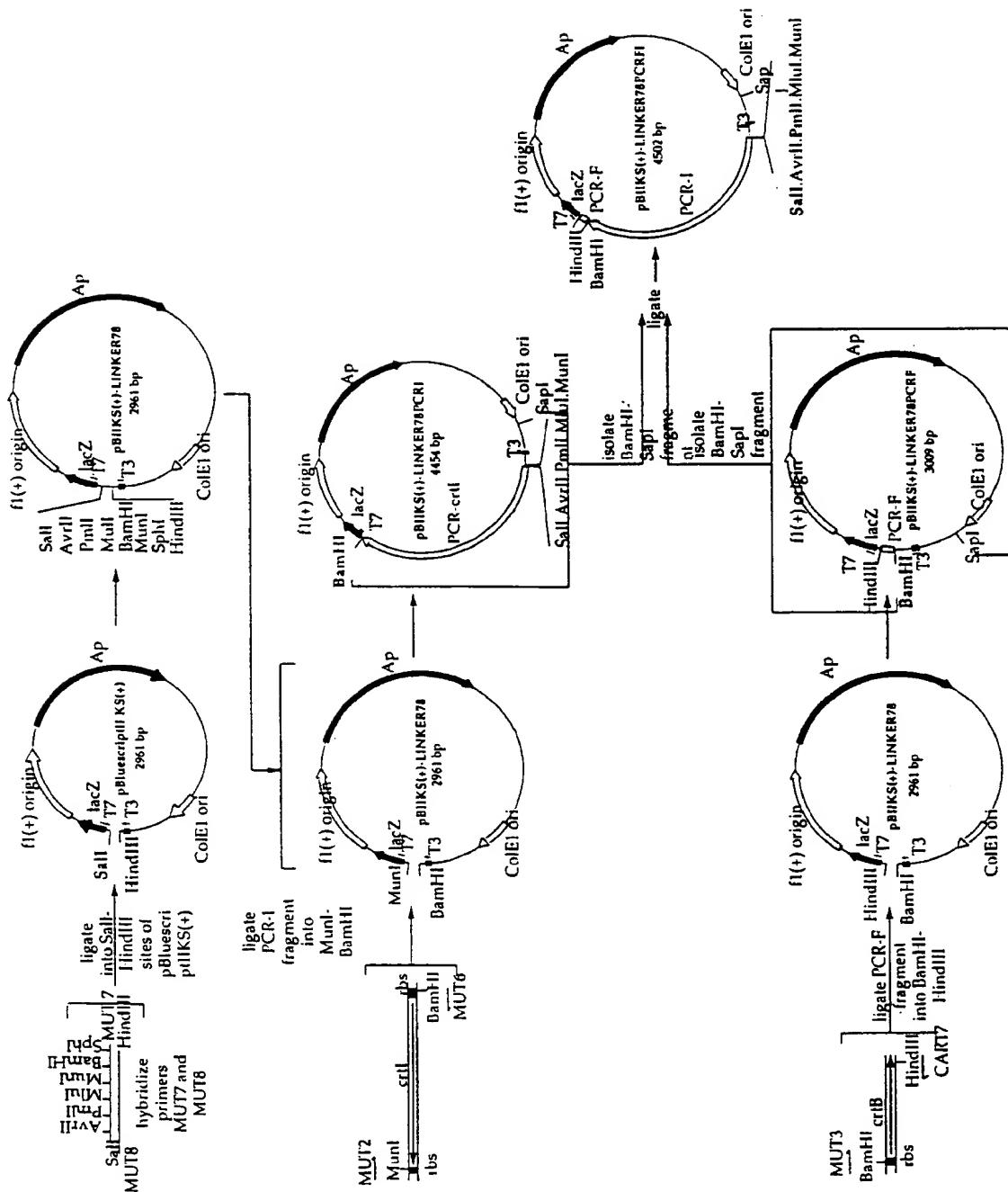


Fig. 20/2

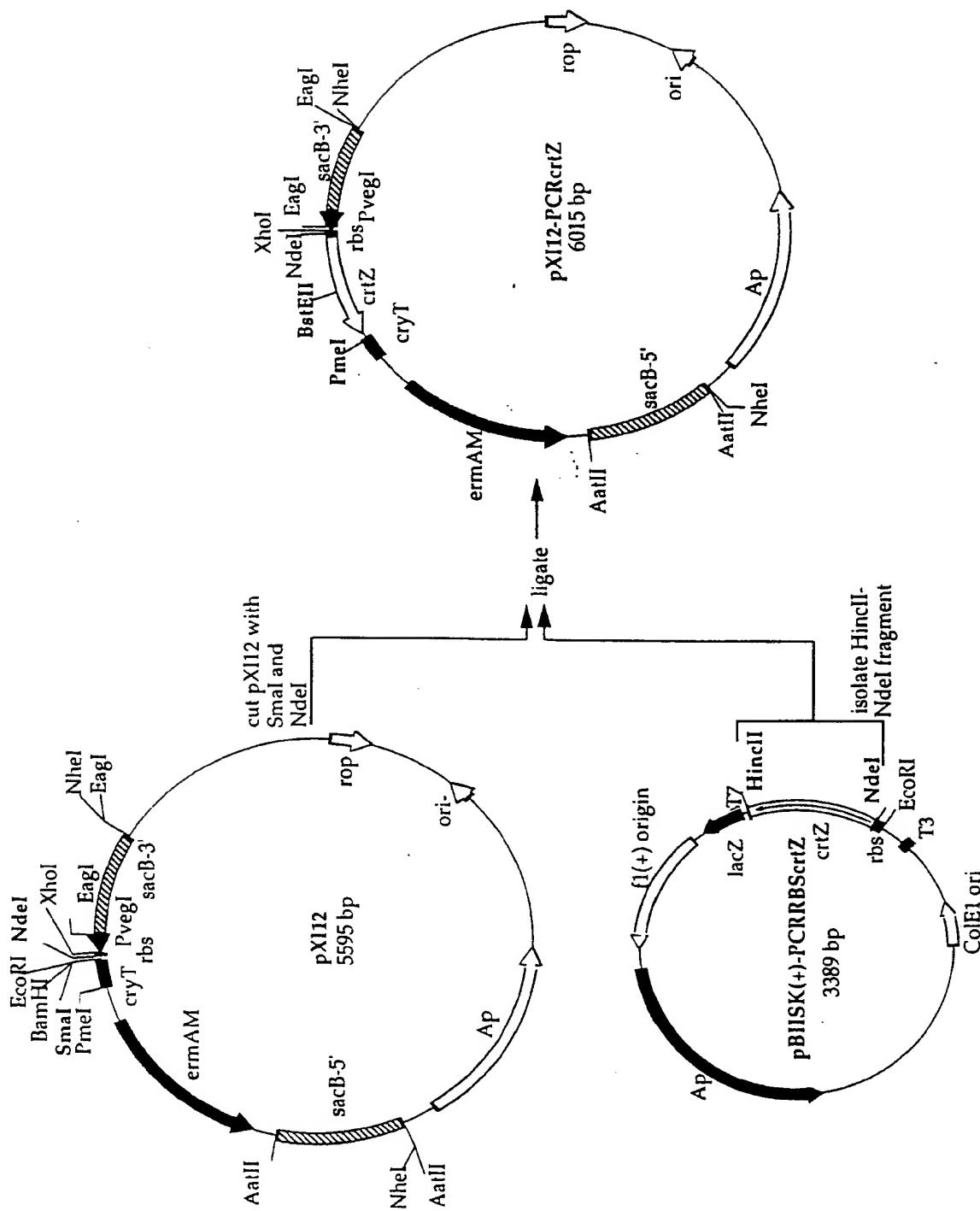


Fig. 20/3

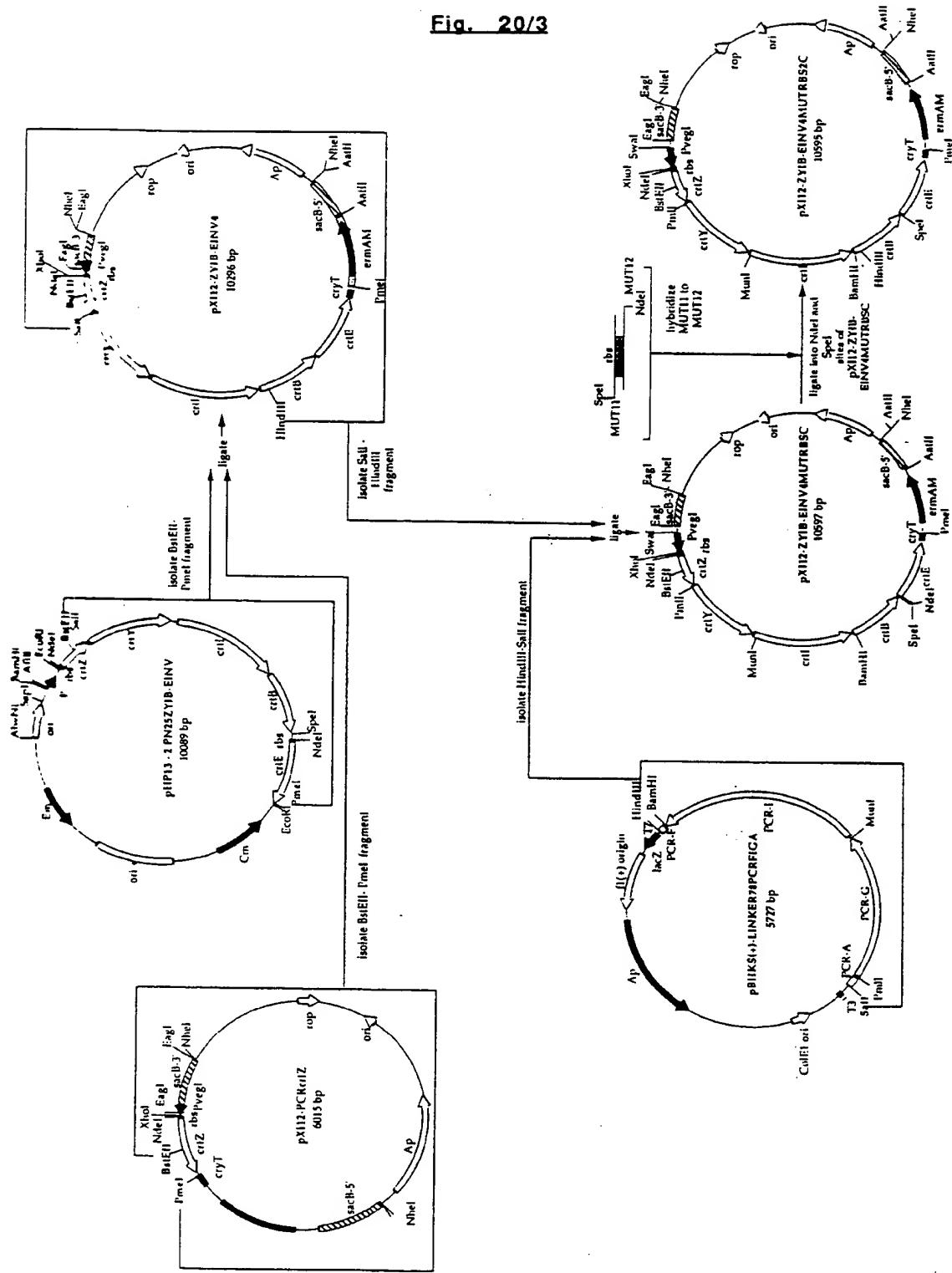


Fig. 20/4

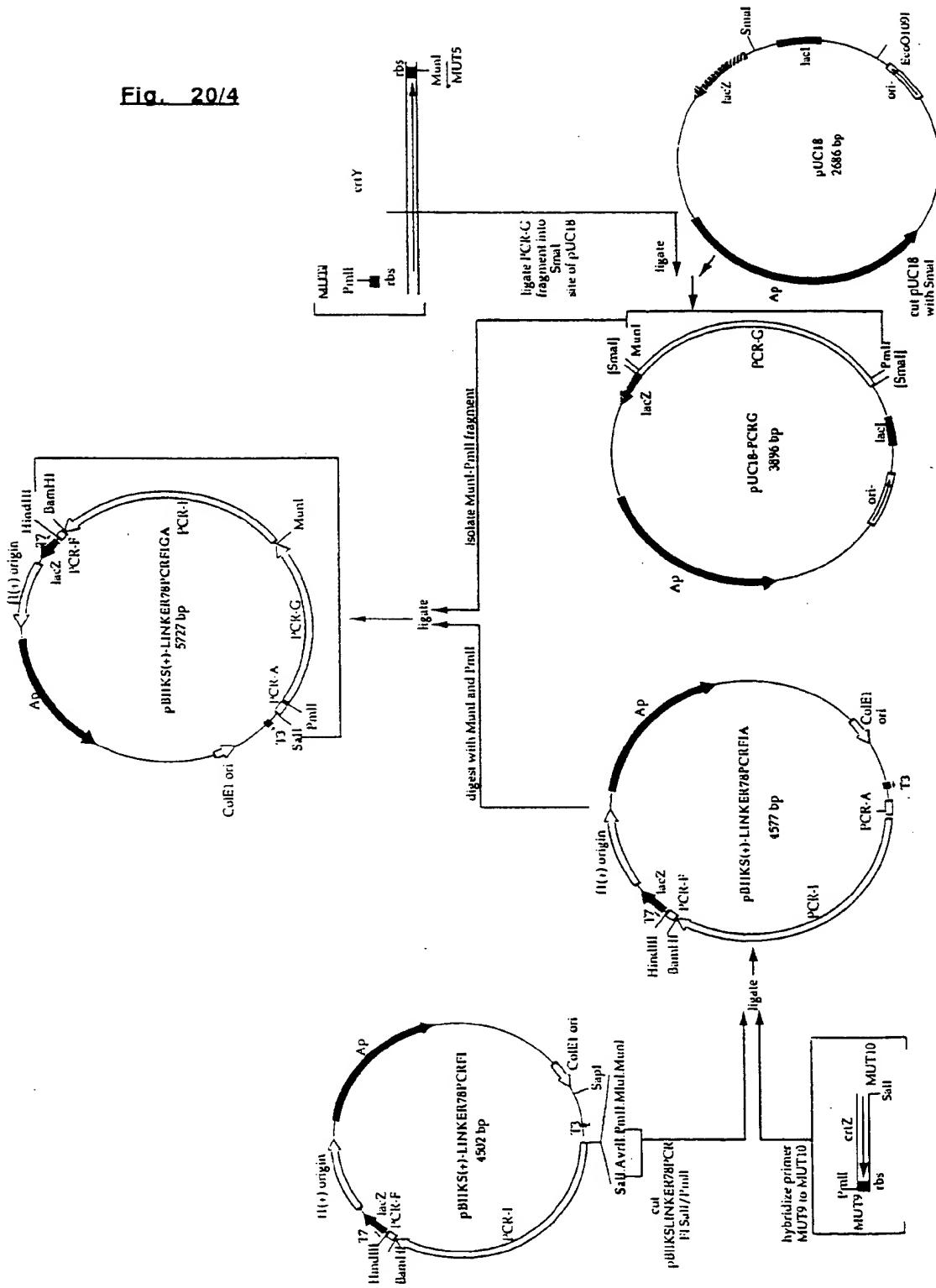


Fig. 21/1

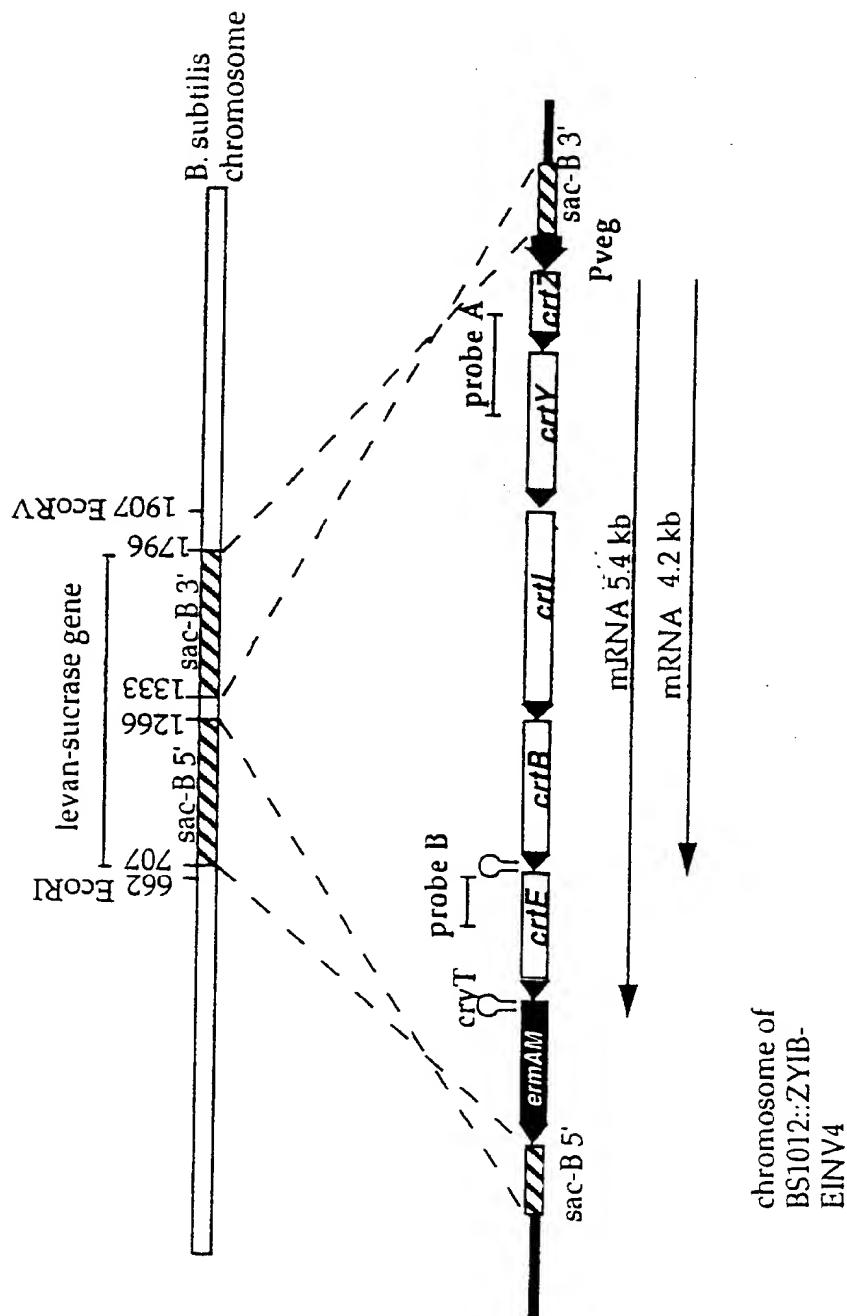


Fig. 21/2

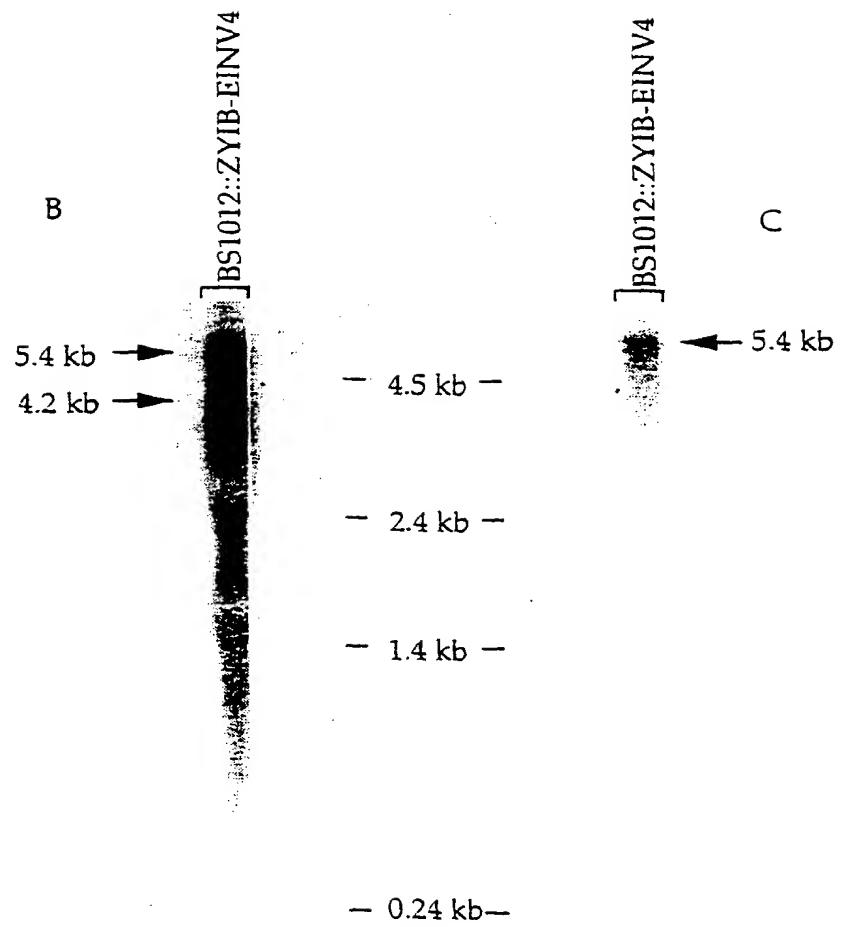


Fig. 22

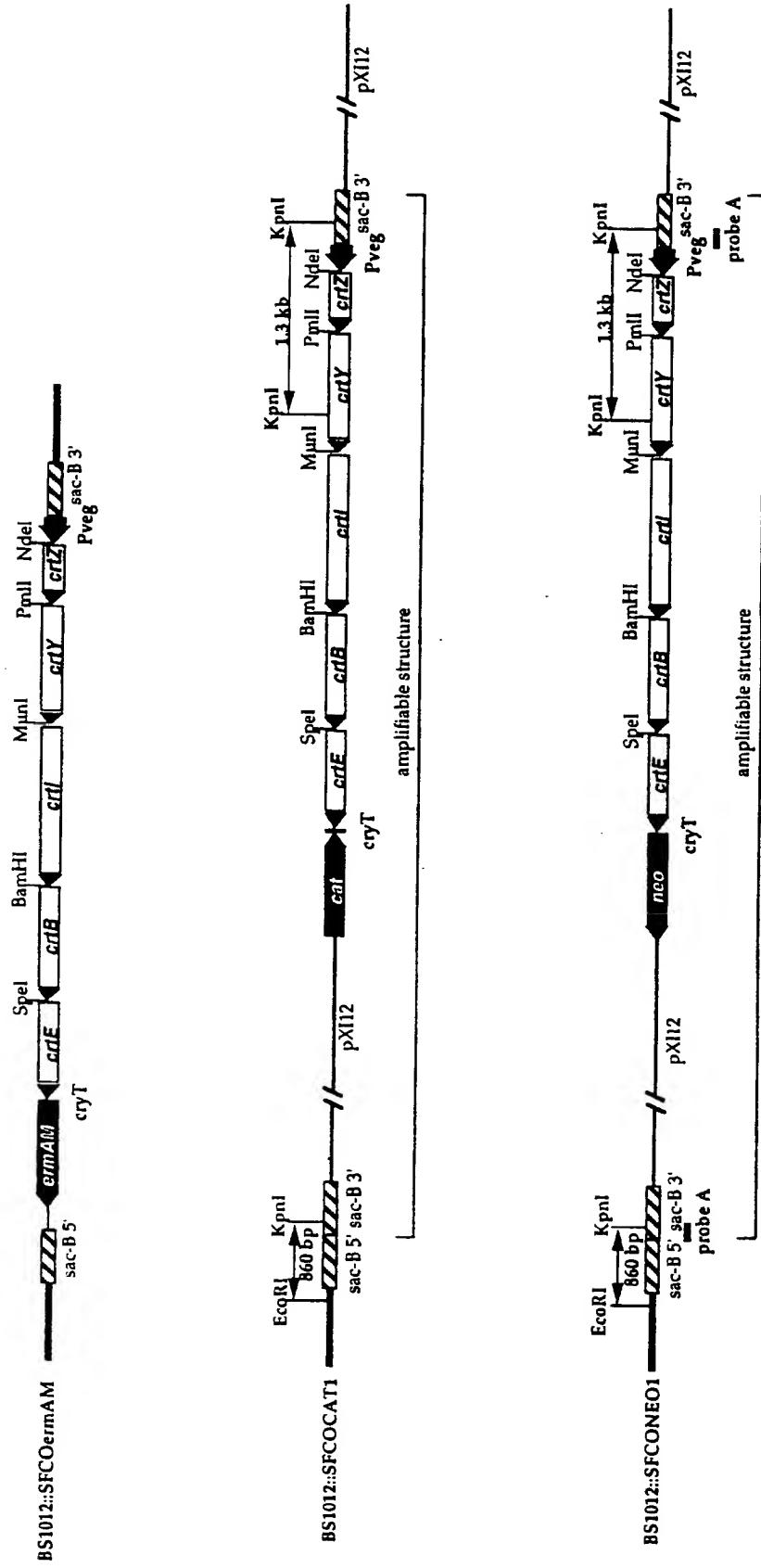


Fig. 23

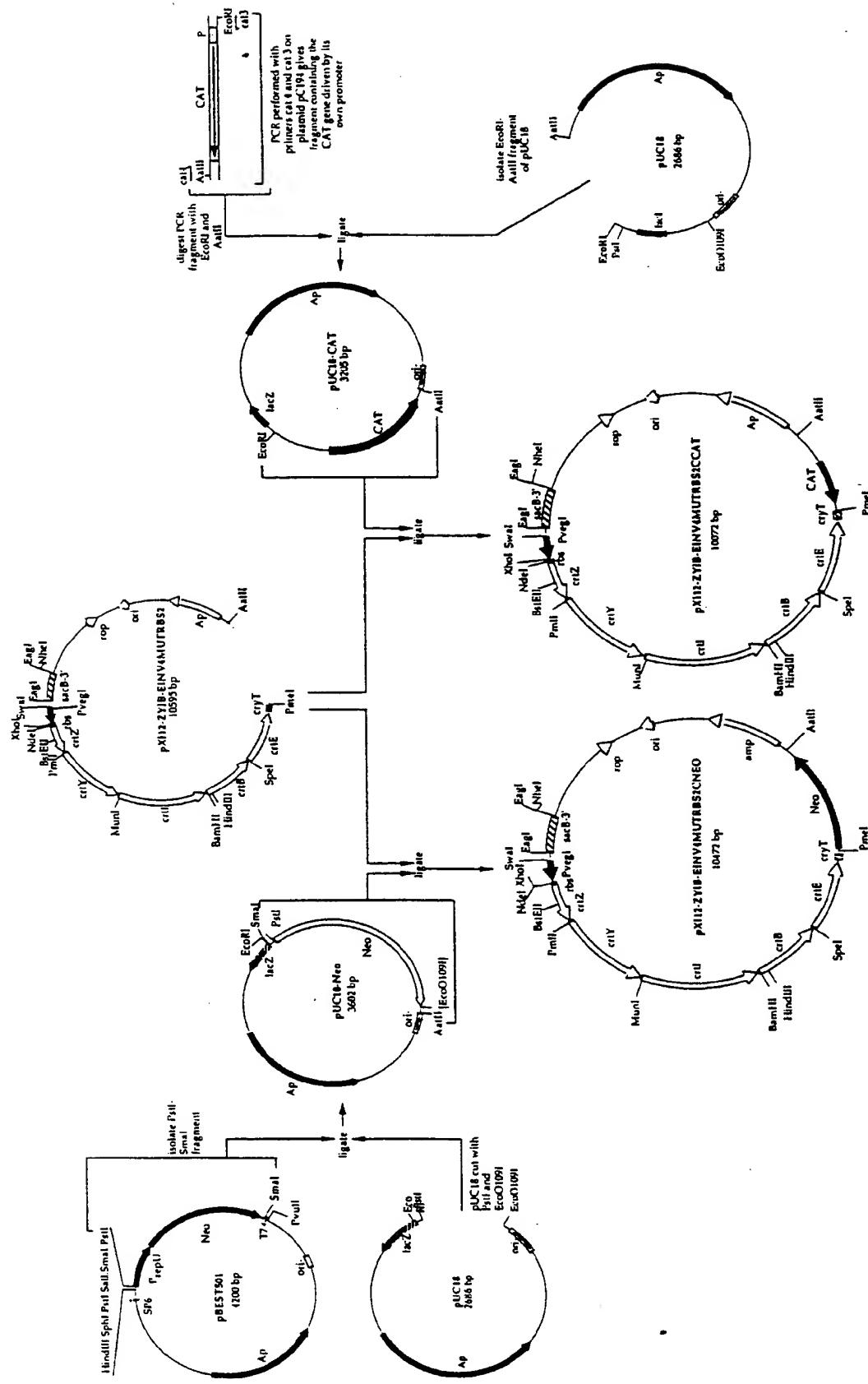


Fig. 24/1

1 CTAATTGTAAGCGTTAATATTTGTTAAATTCCGCGTTAAATTTTGTAAATCAGCTC  
 60  
 GATTTAACATTGCAATTATAAAACAATTAAAGCCAAATTAAACAAATTAGTCGAG  
 61  
 ATTTTTAACCAATAGGCCGAAATCGGAAATCCCTTATAAAATCAAAAGAATAGACCGA  
 120  
 TAAAAAAATTGGTTATCCGGCTTATAGCCGTTAGGGAAATTAGTTCTTATCTGGCT  
 121  
 GATAGGGTTGAGTGTGTTCCAGTTGGAAACAAGAGTCCACTATTAAAGAACGTGGACTC  
 180  
 CTATCCCACACTCACAAACAAGGTCAAACCTTGTCTCAGGTGATAATTCTGCACCTGAG  
 181  
 CAACGTCAAAGGGCGAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACC  
 240  
 GTTGCAGTTCCCGCTTTGGCAGATAGTCCCGTACCCGGTATGCACTTGGTAGTGG  
 241  
 CTAATCAAGTTTTGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG  
 300  
 GATTAGTTCAAAAACCCCAGCTCCACGGCATTCTGATTTAGCCTTGGGATTCCCTC  
 301  
 CCCCCGATTAGAGCTTGACGGGAAAGCCGGGAACGTGGCAGAAAAGGAAGGGAAAGAA  
 360  
 GGGGGCTAAATCTCGAATGCCCTTCCGGCGCTTGACCCGCTTTCTTCCCTCTT  
 361  
 AGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACCGCTGGCGTACAC  
 420  
 TCGCTTCTCGCCCGGATCCCCGACCGTTACATGCCAGTGGGACGCCGATGGTG  
 421  
 CACACCCGCCGCGCTTAATGCCCGCTACAGGGCGCTCCATTGCCATTAGGCTGCC  
 480  
 GTGTGGCGGCGGAATTACGGCGATGTCGGCGAGGGTAAGGGTAAAGTCCGACGC  
 481  
 CAACTGTTGGGAAGGGCGATCGGTGCGGGCTTTCGCTATTACGCCAGCTGGCGAAAGG  
 540  
 GTTGACAACCCCTCCGCTAGCCACGCCGGAGAACGATAATGCCGCGTACCGCTTCC  
 541  
 GGGATGTGCTGCAAGGCCATTAGTTGGTAACGCCAGGGTTTCCAGTCACGACGGTG  
 600  
 CCCTACACGACGTTCCGCTAATTCAACCCATTGGTCCAAAAGGGTCACTGCTGCAAC  
 601  
 TAAAACGACGGCCAGTGAGGCCGTAATACGACTCACTATAGGGCAATTGGAGCTCCA  
 660  
 ATTTTGTGCCGGTACTCGCGCCATTATGCTGAGTGTATCCGCTTAAACCTCGAGGT  
 661  
 CCGCGGTGGCGCCCTCTAGGGATCCGCCCTGGCGTTCGCGATCAGCAGCCCGCC  
 720  
 GGGCCACCGCCGGGAGATCACCTAGGCCGGACCGGAAGCGCTAGTCGTGGCGGGGA  
 721  
 TGCGGATCGGTCAAGCATCCCCATGAACCGCAGCGCACGCGAGCGCGCCAGA  
 780  
 ACGCTAGCCAGTCGTAGTAGGGTACTTGGCGTCCGCTGCGTCCGCGCGGGTCT  
 781  
 TCGGGCCCGTCCAGCACGGCATGCCCATCTCGCAAGGCCCGGCGATGGGGCGC  
 840  
 AGCCCGCCAGGTCGTGCCGTACGCCGGTAGTACGCCCTCCGGGGCGCGTACCCCGCG  
 841  
 GTGCCCATCCGAAGAACTCGCAGCCGTCCGCTGCCAAGGTGGCCAGATCGCGCCG  
 900  
 CACGGGTAGGCTTCTGAGCGTCCGACAGGGCACGCCCTCCAGCGGGTCTAGGCCGGC  
 901  
 TATTCCGATGCACTGACGGGCGGATGCCGTGGGCCCTGCCCGCCGCCACCAGC  
 960  
 ATAAGGCATACGTCACTGCCGGGCTACGCCACCCGGGGAGCGGGCGGGTGGTCG

Fig. 24/2

961 GCATCGCGCACGAACCCCTCCGAGATGATGTGCTGATCCATGGCCCGTCATTGCCAAAACC 1020  
 961 CGTAGCGCGTGTGGAAAGGCTACTACACGACTAGGTACCGGGCAGTAACGTTTTGG  
 1021 GATCACCGATCCTGTCGCGTGTGGCATTGTTGCAATGCCCGAGGGCTAGGATGGCGC 1080  
 1021 CTAGTGGCTAGGACAGCGCACTACCGTAACAAACGTTACGGGCTCCGATCCTACCGCG  
 1081 GAAGGATCAAGGGGGGAGAGACATGGAAATCGAGGGACGGGCTTGTGCGTCACGGGGC 1140  
 1081 CTTCCCTAGTTCCCCCCCCTCTGTACCTTAGCTCCCTGCCAGAAACAGCAGTGGCCGC  
 1141 CCGCATCGGGTCTGGGGCGGCCCTGGCGGGATGCTGGCCCAAGGCGGGCGGAAGGTCG 1200  
 1141 GGCCTAGCCCCAGACCCCCCGCCGGAGCCGCCCTACGACCGGGTCCGCCCGCGCTTCCAGC  
 1201 TGCTGGCCGATCTGGCGGAACCGAAGGACGGCGCCCGAAGGCGGGTTACGGGGCTGCG 1260  
 1201 ACGACCGGCTAGACCGCCCTGGCTTCCGTGCGGGCTTCCGCCAGTGGCCCGGACGC  
 1261 ACGTGACCGACGGACCGCTGGCAGACGGCCATCGCGCTGGCGACCGACCGCTTCCGCA 1320  
 1261 TGCACCTGGCTGCGCTGGCGACGGCTCTGGCTAGCGCGACCGCTGGCTGGCGAGCGT  
 1321 GGCTGGACGGCTTGTGAACTGGCGGGCATGGCGGGCGAACGGATGCTGGCCCGCG 1380  
 1321 CCGACCTGCCGGAACACTTGACCGCCCGTAGCGCGGGCTTGCCTACGACCCGGCGC  
 1381 ACGGGCCGATGGACTGGACAGCTTGGCGGTGCGGTACGGATCAACCTGATGGCAGCT 1440  
 1381 TGCCCCGGCTACCTGACCTGTCGAAACGGGACGCCAGTGGCTAGTTGGACTAGCCGTCG  
 1441 TCAACATGGCCGCCCTGCAAGCCGAGGCCATGGCCGGCTCGCTACCGGGCTTGCCTGGCGAGCGC 1500  
 1441 AGTTGTACCGGGCGGAACGTCGGCTCCGCTACCGGGCTTGCCTGGCGAGGGCCCCGCTCG  
 1501 GTGGCGTATCGTCAAACCGGCTCGATCGCGGCCGAGGCGATGGCCCGAACGGCGTCCGGCGAGC 1560  
 1501 CACCGCACTACCAAGTGTGCGGGAGCTACCGCCGCGTCTGCCTGCTAGCCCTGCCCAGC  
 1561 CCTATGCCGCCAGCAAGCGGGCGTGGCGGGCATGACCGCTGCCATGGCCCGACCCCTTG 1620  
 1561 GGATACCCGGTCTTCCGCCCGCACGCCCGTACTGCGACGGTACCGGGCGCTGGAAC  
 1621 CGCGGCCACGGCATCCGCGTATGACCATCGCGCCGGCATCTTCCCAACCCGATGCTGG 1680  
 1621 GCGCCGTGCCGTAGGCGCAGTACTGGTAGCGCGGGCGTAGAAGGGGTGGGCTACGACC  
 1681 AGGGCTGCCGAGGACGTTCAAGGACAGCCTGGCGGGCGTGGCGGGTGCCTTCCCTCGCGC 1740  
 1681 TCCCCGACGGCGTCTGCAAGTCTGCGACCCGCCGCCACGGGAAGGGGAGCCCG  
 1741 TGGGAGAGCCGTGGAATACCGGGCGTGTGCAACCATCATCGGAACCCCATGCTGA 1800  
 1741 ACCCTCTCGCAGCCTTATGCCCGCGAACACGTCGGTGTAGTAGCGCTGGGTACGACT  
 1801 ACGGAGAGGTATCCGCGTCGACGGCGATTGCGCATGGCCCCAAGTGAAGGGAGCGTT 1860  
 1801 TGCCTCTCCAGTAGGCGGAGCTGCCCGCTAACCGCGTACCGGGGTTCACTTCTCGCAAA  
 1861 CATGGACCCCATCGTCATCACCGGGCGCATGCCGACCCCGATGGGGGCACTCCAGGGCGA 1920  
 1861 GTACCTGGGTAGCAGTAGTGGCGCGCTACGCGTGGGCTACCCCGTAAGGTCCCCT  
 1921 TCTTGCCTGATGGATGCCCGACCCCTGGCGCGGACCGATCCCGCCGCGTGAACGG 1980  
 1921 AGAACGGCGTACCTACGGGCTGGGAACCGCGCCTGCCATGGCGGGCGCACTTGCC

Fig. 24/3

1981 CCTGTCGCCCCGACATGGTGGACGAGGTGCTGATGGGCTGCGTCCCTCGCCGCGGGCCAGGG + 2040  
 GGACAGCGGGCTGTACCACCTGCTCCACGACTACCCGACGCGAGGAGCGGCCGGTCCC  
 TCAGGGACCGGGCACGTCAAGCGGGCTTGGCGCCGACTGCCGCTGTGCGACGGGACGAC + 2041  
 AGTCCGTGGCCGTGAGTCGCGCGAACCAGGGCTGACGGGACAGCTGCCGTGCTG + 2100  
 CACCATCAACGAGATGTGCGGATCGGGCATGAAGGGCGATGCTGGGATGACCTGAT + 2101  
 GTGGTAGTTGCTCTACACGCTAGGCCGTACTTCCGGCCTACGACCCGGTACTGGACTA + 2160  
 CGCCGCGGGATCGGGCATCGTCGCGCCGGATGGAGAGCATGTCGAACGCC + 2161  
 GCGGCGCCCTAGCCGCCGTAGCAGCGCCGCCCTACCTCTCGTACAGCTTGGGGGG + 2220  
 CTACCTGCTGCCAAGGCGCGTGGGATGCGCATGGGATGACCGTGTGGATCA + 2221  
 GATGGACGACGGGTTCCGCGCCAGCCCCCTACCGTACCCGGTACTGGCACCGACCTAGT + 2280  
 CATGTTCCCTGACGGGTTGGAGGACGCCATGACAAGGGCCCTGATGGGACCCCTCGC + 2281  
 GTACAAGGAGCTGCCAACCTCCTGCGGATACTGTTCCCGGGACTACCCGGAGCG + 2340  
 CGAGGATTGCGCCGGGATCACGGTTTACCCGCGGGCGAGGACGACTATGCGCTGAC + 2341  
 CCTCTAACCGGGCGCTAGTGCCAAAGTGGCGCTCCGCGTCTGCTGATAACGCGACTG + 2400  
 CAGCCTGGCCCGCGCGCAGACGCCATGCCAGCGGTGCCCTGGCCGAGATCGGCC + 2460  
 GTCGGACCGGGCGCGTCTGCGGTAGCGGTGCCACGGAACCGGGCGCTTAGCGCGG + 2520  
 CGTGACCGTCACGGCACGCAAGGTGCAAGGACACCACCGTCGATACCGACGAGATGCCGGCAA + 2581  
 GCACGGCAGTGGCTGCGTTCCACGTCGGTGGCAGCTATGGCTGCTCTACGGGCGTT + 2580  
 GGGCCGCCCCGAGAAGATCCCCATCTGAAGCCGCCCTCCGTGACGGTGGCACGGTCAC + 2581  
 CCGGGCGGGCTCTCTAGGGGGTAGACTTCGGGCGAAGGCAGTGCACCGTGCCAGTG + 2640  
 GGGCCGAAACAGCTCGTCGATCTGGACGGGGCGGGCGCTGGTATGATGCGCCAGTC + 2641  
 CCGCCGCTTGTGAGCAGCTAGAGCCTGCCCCGCCCGACCAACTACCGGGTCAG + 2700  
 GCAGGGCGAGAAGCTGGGCTGACGCCATCGCGGGATCATCGGTATGCCGACCCATGC + 2761  
 CGTCCGGCTCTCGACCCGGACTGCGGCTAGCGCGCTAGTAGCCAGTACGCTGGTAGC + 2760  
 CGACCGTCCCCGCTGTTCCGACGGCCCCATCGCGCGATGCGCAAGCTGCTGGACCG + 2761  
 GCTGGCAGGGCGGACAAGGGCTGCCGGGGTAGCCGCGTACCGTTCGACGACCTGGC + 2820  
 CACGGACACCCGCTTGGCGATTACGACCTGTCGAGGTGAACGAGGCATTGCGCTCGT + 2821  
 GTGCCGTGCGGAAACCGCTATGCTGGACAAGCTCCACTTGCTCCGTAAGCGGAGCA + 2880  
 CGCCATGATCCCGATGAAGGGCTTGGCCTGCCACAGATGCCACGAACATCAACGGCG + 2881  
 GCGGTACTAGCGCTACTTCCCGAACCCGACGGTGTGCTACGGTGGCTGATGCGCC + 2940  
 GGCGCTGCGCGCTTGGCCTCCCGATCGCGCGTGGGGCGCGGATCATGGTCACGCTGCT + 2941  
 CGGACGCGCGAACCCGTAGGGTAGCCGCGCAGCCCCCGCGCTAGTACCAAGTGCAGA + 3000  
 GAACGGCATGGCGCGCGGGCGGACGGCGGGGGCGCATCCGTCTGCATCGCGGGGG + 3001  
 CTGCGCTACCGCGCGCCCCCGCTGCGCGCCCCGGTAGGGCAGACGTAACCGGCCCCC + 3001

Fig. 24/4

3001 CGAGGCGACGCCATCGCGCTGGAAAGGCTGAGCTAATTCAATTGCGCGAATCCCGCTTT 3060  
 GCTCCCGTCCGGTAGCGCGACCTTGGCGACTCGATTAAGTAAACCGCGTTAGGCGCAA  
 TTGCGCACGATGGGGAACGGAAACGCCACGCCGTGTTGGCTTGGCTGGACCTGTCT 3061 3120  
 AAGCACGTGCTACCCCCCTTGGCTTGGCGGTGGGACAACACCAACCGCAGCTGGACAGA  
 TCGGGCCATGCCGTGACCGATGTGGCAGCGCATGGGCGTTGCCGATCCGGTCGCAT 3121 3180  
 AGCCCCGTACGGGCACTGCCTACACCGTCCCGTACCCCGAACGGCTAGGCCAGCGTA  
 GACTGACGCAACGAAGGCACCGATGACGCCAACGAGCAATTCCCCCTACCGATCTGGT 3181 3240  
 CTGACTGCGTTGCTTCCGTGGCTACTGCGGGTTCTGCGTTAAGGGGGATGCGCTAGACCA  
 CGAGATCAGGCTGGCGCAGATCTCGGCCAGTCCGGCTGGCTCGGCCCCGCTGGCGC 3241 3300  
 GCTCTAGTCCGACCGCGTCTAGAGCCCCGTCAGCCCCAACCGCTTCCGGCGTGTGATGCTGAT  
 GGCCATGAGCGATGCCGCCCTGTCCCCCGGCAAACGCTTCCGGCGTGTGATGCTGAT 3301 3360  
 CCGGTACTCGCTACGGCGGGACAGGGGCGTTTGCAGAAAGCGCGGACGACTACGACTA  
 GGTGCGCAAAGCTCGGGGGGTCTGCGATGCGATGGCTGATGCCGCCCTGGCGGTGCA 3361 3420  
 CCAGCGGCTTCTGAGCCCCCCCCAGACGCTACGCTACCGCTACGGCGAACGCCAGCT  
 GATGGTCCATGCCGATCGCTGATCTCGACGACATGCCCTGATGGACGATGCCAGGAC 3421 3480  
 CTACCAAGGTACGGCGTAGCGACTAGAAGCTGCTGTACGGGACGTACCTGCTACGGTCTG  
 CCGTCGCGGTCTGCCGCCACCCATGTCGCCCATGGCGAGGGGCGCGCGGTGCTGGGG 3481 3540  
 GGCAAGGCCAGTCGGCGGTGGTACAGCGGTACCGCTCCCCCGCGGCCACGAACGCC  
 CATGCCCTGATCACCGAGGCCATGGGATTTGGCGAGGCCGCCGCCAGGCCGGGA 3541 3600  
 GTAGCGGGACTAGTGGCTCCGGTACGCTAAACCCGCTCCGGCGGCCGCGCTGCGGCC  
 TCAGCGCGAAGGCTGGTCGATCCATGTCGCCCGCATGGGACCGGTGGGCTGTGCGC 3601 3660  
 AGTCGCGCTTCCGACCGCGTAGGTACAGCGCGCGTACCCCTGGCCACCCGACACCGG  
 AGGGCAGGATCTGGACCTGACGCCCAAGGACGCCGGATCGAACGTGAAACAGGA 3661 3720  
 TCCCGTCTAGACCTGGACGTGCGGGGTTCTGCGGCCCTAGCTGCACTTGTCT  
 CCTCAAGACCGCGGTGCTTCTGCGGGCCTCGAGATGCTGCAATTATAAGGGTCT 3721 3780  
 GGAGTTCTGGCCGACGACAAGCAGGCCGGAGCTACGACAGGTAAATAATTCCAGA  
 GGACAAGGCCAGACCGAGCAGCTATGGCCTTGGCGTACCTGGTGGCTTCCA 3781 3840  
 CCTGTTCCGGCTGGCTCGCAGTACCGGAAGCCCGCAGTCGAACCCAGGCCAGAGT  
 GTCTATGACGACCTGCTGGACGTGATCGCGACAGGCCAGCACCGGCAAGGATACGGC 3841 3900  
 CAGGATACTGCTGGACGACCTGCACTAGCCGCTTCCGGTGGCGTCTATGCCG  
 GCGCGACACCGCCGCCCGGCAAGGGCGGCTGATGGCGGCGGACAGATGGCGA 3901 3960  
 CGCGCTGTGGCGGGCGGGGTTCCCGCCGGACTACCGCCAGCCGTCTACCCGCT  
 CGTGGCCAGCATTACCGGCCAGCCGCGCACTGGACGAGCTGATGCGCACCCGGCT 3961 4020  
 GCACCGCGTGTGACGACCTGCACTAGCCGCTTCCGGTGGCGTCTATGCCG

Fig. 24/5

4021 GTTCCGGGGGGGAGATCGCGGACCTGCTGGCCCGGTGCTGCCGCATGACATCCGCCG 4080  
 CAAGGCGCCCGTCTAGCGCTGGACGACGGCGACGGCGTACTGTAGGCGGC  
  
 4081 CAGCGCTAGGCGCCGGTCGGTCCACAGGCCGTCCGGCTGATTCGCCGCCAG 4140  
 GTCGCGGATCCGCGCCAGCCAGGTGTCGGCAGGCCGACTAAAGCGGCGCGTC  
  
 4141 GCGCGATGCGGCCCGTCCAAGCCTCCGCCAGAAGCCGATTTGGCAGCCTTCGA 4200  
 CGCGCTACGCCGGCGCAGGTTCGGAGGCCGGTCTTCGGCTAGAACCGTCGGAAGCT  
  
 4201 CGTGCATCCGCTGGCGATAGGCCCTCGGGCCACCCGTCCGGATGCGGTCCCATTGC 4260  
 GCACGACTAGGCACCGTATCCGAGCCCGGTGGGACGGCTACGCCGAGGGCTAACG  
  
 4261 GCGATAGATAACGACGCCGGCGATCGACCAACGCCAGCGCGGGAGATGCGGAAG 4320  
 CGCTATCTATGCGTCGCCGCCGCTAGCTGGTGCCTCGCCGCCGCTACGCCCTTC  
  
 4321 CCCCTGCCGCCGCCAGGCATAATAGGCCCTGGCCGCCGCTAACGAGGCCGATGATGCGGA 4380  
 GGGGACGCCGCCGCCCTCCGTATTATCCCGAGCCGGCGTAGTCGCTCCGCTACTACTGCC  
  
 4381 ATAGAGCGCTCCGAGGCCACCCGACCTCAACCGTCGCCCGCTCGCCAGCCAGTC 4440  
 TATCTCGCCAGGCCCTCCGTGGCCTGGAGTTGGCAGCGGGGGCGAGCCGGTCGGTCAG  
  
 4441 GGCAGGCAGATAGCAGCGCCGATGGCGGATCGCGATCACGTCGAGCGATGTTCGT 4500  
 CCGTCCGTCTATCGTCGCCGCTACCGCCCGTAGCAGCTAGTGCAGCGCTCGCTACAAGCA  
  
 4501 CAGCTGGAACGCAAGGCCAGATCGCAGGCCGATCCAGCACCGCATCGCCTGCCACGCC 4560  
 GTCGACCTTGCCTCCGGCTAGCGTCCCGCTAGTCGTTGGCTAGCAGGACGTGGG  
  
 4561 CATCACCCGCCATCATCACGCCACGACCCCCCGGACGTGGTAGGAATATTCCAGCAC 4620  
 GTAGTGGCGCGGTAGTAGTGGGGTCTGGGGCGCTGACCATCCTTATAAGGTCTG  
  
 4621 GTCATCCAGGCTGGGTATTCGGCATCCGCCACATCCATCGGAAACCCCTCGATCAGGTC 4680  
 CAGTAGGTCCGACGCCATAAGCGCTAGGCCTGAGTAGCGCTTGGGAGCTAGTCCAG  
  
 4681 CATCGGCCAAAGTCCGGAAATCATGCCGCCGGGACCTGGCGAGCGCCGCCAGGG 4740  
 GTAGCCGGTTCCAGGCCCTTAGTACGGGGCCGCTGGACCGCGTCGCCGCCCTCCC  
  
 4741 CGCGACATCGGGCGTCTCGTCAGCGCGGCCAGCGTGTGGCGCGAGCGCCCCCAG 4800  
 GCCGCTGTAGCCGGCAGGAGCACGTGCGCCGGTCGACAGCCGCCGTCGCCGGGGTC  
  
 4801 CGCGCCCTGTGGTCGCCGCCCTCGGGGGAGAACCCATCACCTGCCGTGATCAC 4860  
 CGCGCGACACCCAGCGGGGGGGAGCCCCCGTCTGGTAGTGGACGGCAGCTAGTG  
  
 4861 GTCATCCGATGCCATGCCGACAGGCATAGAGCATGACCGTATCCCGGGATGCCGGGG 4920  
 CAGTAGGCCTACGGACGTGGTCCGTACTGGCATAGGAGGCCCTACGGCCCGCC  
  
 4921 CATCAGCTGGCCGCGTGCAGCAAGCTTGGCAACCCCTGCCGATGCCGCTTCGGAAAGT 4980  
 GTAGTCGAACCGGGGACGCGCTTCGAAACGCTTGGGACCGCGTACCCGCCAAGCCTTC  
  
 4981 CGCCGTCAGATCGGTATGCCACGGCCAGGTCCGACAGCATGACCTGCGCCGTGGCTTG 5040  
 GCGGCAGTCTAGCCAGTACGCTGCCGGTCCAGGCTGCTGACTGGACGGCCACCGGAAC

Fig. 24/6

5041 GCGCTGCCAACGACACCCGGATGCCCGACCCGGATGCCGTGCCGCCACGATGTAG 5100  
 CGCGACGGTTGCTGTGGGCCCTACGGCGTGGCCCTACGCACGGCGGGGTGCTACATC  
 5101 AAGTTCGGGATCGCGGGTCCGGTTATGCCGGCGAACCAGGGGATTGCCGTAGGATC 5160  
 TTCAAGCCCTAGCGGCCAGGCCAATACGCCGCCCTGGTCGCCCTAACCGAGTCCTAG  
 GGCTCGACCGAGAAGGCGCTGCCGTGATGGGCCACAGTCGGTGTGAAATCGGGGGG  
 5161 CCGAGCTGGCTTCCCGACGGCACTACCCGGCTGTCAAGCCACGACTTAGCCGCC 5220  
 CTGAAGATGCGGCTGACGGTCAGGTGCTTGCAGGGTGGGATGGCGGGCTCCAGT  
 5221 GACTTCTACGCCACTGCCAGTCCACCGAACCGTCCAGCCCTACCGGCCGGAGGTCA 5280  
 TCCTCGAAGATGCGCTCGCATGCCCGGGCTCGGTTCCAAATGACATCGGGCGGG  
 5281 AGGAGCTTCTACCGAGCCGTATCGGCCCCGGAGCCGAGGGTTAGCTGTAGCCGCC 5340  
 CCCAGATGCCGAACGGGCCAACGGACGTAATGCGTGGACATCCCTCGGGGCCAGGCTG  
 5341 GGGTCTACGCCCTGCCCGTTCCTGCATTACGACCTGTAGGGAGCCCCGGTCCGAC 5400  
 GGATCGGTACGCAGGGCAATGCAGATACTCGAGAAATCGTCCGGCAGGGTGGCCCG  
 5401 CCTAGCCAGTGCCTCCGTTACGTCTATGTAGCTCTTAGCAGGCCGTCCGACCGGGC 5460  
 TTGAAGATCTCGTCAACGCCCTTGATAGCGCGGGCGAACGGATGACGCTGTGGTGGCC  
 5461 AACTCTAGAGCAAGTGGTGGGAACATCGGCCGGCTCTACTGCGACACCAACCGG 5520  
 AGTTCTCGGGCGCTTGGACAGGCCAACATGCAACAGCGACATCGACCAAGCGC  
 5521 TCCAAGAGCCCCGCCAACCTGTCGGCTTACGTCGTGCTGCGCTGTAGCTGGTCGG 5580  
 TGCCGGTTCAAGGATCGCGCCTTGGTGCGCCCGCGGGTATGGCCAGCAGGTGGCGA  
 5581 ACGGCCAAGTCTAGCGCCGGAACACCGCGGGCGCCATACCGGGTGTCCAGCGCT 5640  
 TAGCTGTGCATCACGTCGCCGTTGCTGGCACCGTATCCGCGCGCAACTGCCGCCGTCC  
 5641 ATCGACACGTAGTGCAGCGCAACGACCGGTGGCATAGGCGCGCTGACGGCGGGCAGG 5700  
 AGCAGCGTACGCCCGTGGCGATGCCCTCGGTGTCATCCCGTGACGGGGCATTC  
 5701 TCGTCGCACTGCGGGCACCGCGTAGCGGGAGCCACAGCTAGGCGACTGCGCCCGTAAAG 5760  
 AGCAGCAGCGTGCCGCCAACAGCGCTGAACAGGGGACCATGCCCGGACAGCTGGTG  
 5761 TCGTCGTCGACGGCGGTTCTGCGAGCTTGCTCCCGCTGGTACGGGGCTGGTCAAC 5820  
 GTGCCGCCCTGGCAACAGACGCCCGCGCCGTTCCAGCGCATGGATCAGCGCATAG  
 5821 CACGGCGGGAACCGCTTGGCTGCCGGCGCGGCAAGGTCGCGTACCTAGTCCGCTATC 5880  
 ATCGAGCTGGTCGAAACCGGTTCCCGCCGACCAAGCAGCGTGTGAAACGAGAAGGCTGC  
 5881 TAGCTGACCAAGCTTGGCCAGGGCGCTGGTCGACACCTTGCTCTCCGGACG 5940  
 CGCAGATGCCGGTCTGGATGAAGCGGCCACCATGCTGTGGACCGAGGGTATGCCGT  
 5941 GCGTCTACGCCAGGACCTACTTCGCGCGTGGTAGCACACCTGGCTGCCATACGGACG 6000  
 AGGCGCATCGCGCCGGCGCGCTTCAGCATCTGGCCAGCTCAGGAAGGGCGTGGTC  
 6001 TCCCGCTAGTCGCGGGCGCGCGCAGCTCGTAGACCGGGTCCGAAGTCCCTCCCGCAC 6060

Fig. 24/7

6061 CCCAGCTTCAGATACCCCTCGCGATAGACCTCCCGCGTAATCGTGGAAAGCGGGGATAG 6120  
 GGGTCGAAGTCTATGGGGAGCGCTATCTGGAGGAGCCGCATTAGCACCTCGCCGCTATC  
 CCATCGACATCGGGGGATTGAAGGGAGGGACCTGGCGGATCAGCTCGTCGTCGTT 6180  
 GGTAGCTGTAGCCGCCCTAACCTCCTCCGCTGGACCGCTAGTCGAGCAGCAGCAAG  
 ACGTATTGAGCTGCGGCCGTCGGCCCATGTCAGCCGGTAGAAGGGGAGACCGGCAGC 6240  
 TGCATAAGCTTCGACGCCGGCAGGGGGTACAGTCGCCATCTTCCCGCTGGCCGTCG  
 AGCGTCACGTCACTGGTGGCCGCTGAGGGCCCACAGCTCTCGCAGGCTGTCG 6300  
 TCGCAGTGCAGTGCAGGGTACGCCAACCGGGACTCCCCGGTGTGAGAGCGTCCGACAGC  
 GCGTCGGTCACGACCGTCGGGCTGCATCGAAGACGTGGCCCTGATCGTCCAGACATAG 6360  
 CCCAGCCAGTGTGGCAGCCCCGACGTAGCTCTGACCCGGACTAGCAAGGTCTGIAIC  
 CGCCGGCCGCCGGGCTTGTGCGGGCCTCGACGATGGTGGTCGGATGCCGGGATTGC 6420  
 CGCGCCGGCGGCCAACAGCGCCCGGAGCTGCTACCCACCGCGCTACGGCCGGCTAACG  
 AGCGGGATGGCAAGCGCAAGCCCGCCGAAACCTGCGCCGATGACCGATGGCGGAACTCATG 6480  
 TCCGCCTACCGTTCGCGTTGGCGGCTTGGACGCGGCTACTGCTACCGCCTTGAGTAC  
 CTCTCTCCTGCAGGGGGCTTGGCAGGCAGCCACGGCCTGCGACAGCGGAATGG 6540  
 GAGAGAGGACGTCGTCGGGCAAGCCGCTCGGTACAGTCCGGGGCGTATCGCTTACCG  
 CGGGCGTCCGGTACGATGCGAAGCCGGTGGCCAATGTCAGGGCCGGCATAGAAC 6600  
 CGCCCCGAGGCCACTGCTACGCTTCGCGCAAGCCGTTACAGTCCGGGGCGTATCGCTTACCG  
 GCTCGATAGCGGCTGGCAGGGTAGAACCGCTGCAAGCAGGGGATAGCGACGGTGG 6660  
 CGAGCTAGTCGCGACGCCGTCGCCATCTGGCAGCTCGTCCGCTATCGCTGCCAGCC  
 CGGGGAGCCGGAACAGCATCCGGTTCAAGCAGCCGAGGAAGCGGTGCGATCCGGC 6720  
 CGCCCCGTCGGCGCCTTGTGCTAGGCCAGTCGTCGGCTCTCGCCAGCGCTAGGCGCG  
 GATCGATGGCCAGCCGGCACCGCGACGGCGGACGGCGTGTCAAGGTGGCGCCG 6780  
 CTAGCTACCGGGTCCGCGCTGGCGCTGCCGCTGCCAGCACTCCAGCGCGCC  
 CGATGGCATCCGGACCTGCGGGCATAGGGCAGCGAATATCCGGTGACGGGGTGGAAACA 6840  
 GCTACCGTAGGCGCTGGACGCCCGTATCCGTCGTTAGGCCACTGCCAACCTTGT  
 GCGCTGCCCGAGCCCAACCGGACCGCCCTGCGCTGGTCGGCCAGAAGCCTATGG 6900  
 CGGGACGGGGTGGGGTGGCGCTGGCGCTGCCGCTGCCAGCAAGCCGGTCTCGGATACC  
 CGTCATGGCCAGCCGATGGCAGGGATGCCCTTCCGCGCCATCTCCGCGGTCC 6960  
 GCAGTACCCGGTCCGCTACCCGTCCTACGGGGAAAGCGCGGCGTAGAGGACGGGCCAGG  
 AGCCCCGCTGGCGCATAGTCAGCGACCCCTGCGCCAGCGCGCCATCGTCCAGATCGC 7020  
 TCGGGGCGGACGCCGTACAGTCGCTGGACGCCGGTGCAGCGCTAGAGGCTAGCG

Fig. 24/8

CGCCGTCGCTGTAGCGCGTATCCTCGATCAGGATGCGGGTGGGACTGAAGGGCAGCAGAT 7080  
 7021 GCGGCAGCGACATCGCGCATAGGAGCTAGTCTACGCCACCCCTGACTTCCCGTCTA  
  
 AGATGAAGCGGTACCCGTCATCTGCGAACGGTCGCGTCCATGATCATGGCGCTCGA 7140  
 7081 TCTACTTCGCCATGGCAGGTAGACGCCCTGCCAGCGCAGGTACTAGTAGCCCGGAGCT  
  
 CGCCATGGGGGCGTGGTCTCGATCTCGACGCCACGAATTCTGAAACCCACGGTCA 7200  
 7141 GCGGTACCCCCCGCAGCCAGAGCTAGAGCTGCGGGTCTAAAGACCTTGGGTGCCAGT  
  
 GGTGCGGGTCTCGACGGCACCAACGGCGTCGATCACCGAGGCAGCTCGATCCGCAGC 7260  
 7201 CCACGCCCAAGAGCTGCCGTGGTGCCCGCAGCTAGTGCCTCCGAGCTAGGCCTCG  
  
 CGTCCGTCAGCGTCGCCGGTATCGTCCAGCGTCGCCACATGCGTATTCCACCGCAGAT 7320  
 7261 GCAGGCAGTCGAGCGCGGCCATAGCAGGTCGAGCGCTGTACGCATAAGGTGGCGTCTA  
  
 CGACACCCCTGCAGCAGCCGATCACGCCGCCCTCGATCGAGCCATAGCCTGCGTCA 7380  
 7321 GCTGTGGGACGTCTCGGGCTAGTCGCCGGGGAGCTAGCTCGTATCGGACAGCAGT  
  
 GGCGGCGCGAATGGTGGGAAACCGCACCTCCTGATCCGTCATTGCCCGACGAATGG 7440  
 7381 CCGCCGGCTTACCAAGCCCTTGCCTGGAGGACTAGGCAGGTAAGCGGGCTGCTTAC  
  
 GCGACAGGCGCGCCAGCCATTGGCGAAAGATCCGTGTCGTGGCAGGACCAGGTGTGCT 7500  
 7441 CGCTGTCCCGCGGGTCGGTAAGCCCGCTTCTAGGCCACAGCACCGTCTGGTCCACACGA  
  
 GGTCCGAGGGCCGGACCGCGCTCGAGCATACGATGCGCGCATCCGGTCTGGTCCGC 7560  
 7501 CCAGGCTCCCCGGCCTGGCGCCAGCTGTAGTGCCTACGCCCGTAGGCCAGACGCCAGCG  
  
 GAACGGCAAGCGCGATCACGCCACCGGACAGCCCCGCCGCGATCAGCAGATCATGGC 7620  
 7561 CTTGCCGTTCGCGCTAGTCGCCGTGGCTGCGGGCGCTAGTCGTCTAGTACCG  
  
 TCATGTATTGGGATCCGCCCCCTCGGGTCTTCAGCAGCGCCCGAGCGTTTCAGCTC 7680  
 7621 AGTACATAACGCTAGGGGGAAAGCGCCAGGAAGTCGTCGCGCGGGCTCGCAAAGTCGAG  
  
 TGCCCTGAGGCTGCGACCGAGGGCGCCAGATGAAACCGAAGCTGACGCAGTTCTCGCG 7740  
 7681 ACGGAACCTCGACAGCTGGCTCCCGCGGGCTACTTGGCTTCGACTGCGTCAAGAGCGC  
  
 GCCATGGACCGCGTGATGCATCCTGTGTGGCTGGTAGACGCCAGGAAGATGCCCGCCT 7800  
 7741 CGGTACCTGGCGCACTACGTAGGACACACGGACCATCTGCCGTGCTCTATCGCGCGAA  
  
 GGGGACATAGCGGAACGCCAGCGCCATGCACCAAGCCGTATGCAGGAATAGTAGAT 7860  
 7801 CCCCTGTAICGCCCTGCCGGTCCGGTACGTGGTCCGAGTACGTCTTATCATCTA  
  
 CAGCCCCGTAGCAGGTGACCCCCACCGCCAGCCACCAGGCCAGATCCGACCCCATCGCGCC 7920  
 7861 GTCGGGCATCGTCCACTGGGGTGGCGGTCCGGTCTAGGCTGGGTAGCGCGGG  
  
 GATCGCGAACAGCACGATCGAGATTACCGCAAGATGACGCCATAGAGGTGCTTCTCTC 7980  
 7921 CTAGCGCTTGTGCTAGCTTAATGGCGCTCTACTGCCGTATCTCCAGCAAGAAGAG

Fig. 24/9

7981 GAGCGCGTGGTCGTATCCTCGTGGTGGCGATTTATGCCAGCCCCAGCCCAGGGGGCC 8040  
 CTCGCGCACCGACACTAGGAGCAGCACCGCTAAATACCGTCGGGTGGGTCCCCCGG  
 8041 ATGCATGATCCACCGATGGACGGAGTAGGCCGTAGCTCCATCGGGCGACGGTCAGGAT 8100  
 TACGTACTAGGTGGCTACCTGCCTCATCGGGCAGTCGAGGTAGGCCGTGCCAGTCCTA  
 8101 GACGGTCAGGATTGGGGCCAAAGTGCTCATGCCGGCCCTTGCTTGATATGACAGGAAAC 8160  
 CTGCCAGTCCTAACGGGGTTACGGAGTACGGCCGGGAACGAACATACTGTCCCTTG  
 8161 AGGCTACGCTGCGCGGGTGCATGACCAGCCCATCGGGGTGCGACCAAAGGGCATCGCG 8220  
 TCCGATGCGACGGCGGCCACGTACTGGTCGGGTAGCCCCACGCTGGTTCCCGTAGCGC  
 8221 TGACATCTGGTTCAGGGCTCATAGGCGATCATCCGTGACATTGCCGCCAACGCGGC 8280  
 ACTGTAGACCCAAGTCCCAGTATCCGCTAGTAGGCACTGTAAGGGGGCTTACCAAGCCCCGC  
 8281 AGGCGCATCACGGGTTCCGTGGTGGAAATATTAATGTTTCCCGAAGATGGTCGGGGCG 8340  
 TCCGCGTAGTGGCAAGCAGGACCTTATAATTACAAAAGGGCTTACCAAGCCCCGC  
 8341 AGAGGATTGCAACCTCCGACCTACGGTACCCAAAACCGTGGCGTACCAAGGCTGGCTAC 8400  
 TCTCTTAAGCTGGAGGCTGGATGCCATGGGTTTGGCAGCGCGATGGTCCGACGCGATG  
 8401 GCCCCGACTGGGAAGGCTTAGCCGATTGTTCCGGCAAGGGAAAGACCTAGTCCGAGGC 8460  
 CGGGGCTGACGCCCTCCGAAATCGGCTAACAGGCCGTTCCCTCTGGATCAGCGTCCG  
 8461 CAGGACCGCATTGCGCCATGCCGGATGCCCATCGGTGACCGGGCTTCAGGCCAAG 8520  
 GTCCCTGGCGTAACAGCGGGTACGGGCTACGGGTAGCCGACTGGCCGAAGTCCGGTTC  
 8521 GCGATCCGCTCTCCGCCCCGATTCGAGGACGAACAGCCGGTGGGGTCCGGATGCC 8580  
 CGCTAGGGCGAGAGGCCGGCGCTAAAGCTCTGGTGTGGCCAGCCCCAGGCCCTAGCGG  
 8581 GACCGCCGCCCCGAATGGGGCTCGTCCAGCGGGCGCATTGGGTGGATGTGGCG 8640  
 CTGGCGGCCGGGCTTACCGCAGAGCAGGTGCGCCGCGTAACGCCACCTACACCGC  
 8641 GATGACGCCGGTTCATCCGAAAGACCATGTCAGCCAGGGATCAGTGTGTTGCCATCCA 8700  
 CTACTGCGGCCAAAGTAGGCCTTCTGGTACAGGTGCCCTAGTCACACAACGGTAGGT  
 8701 GAAGGACACCGGCTGGGGGATTCGTAGATGAAACAGCATCCGGTCCCCCAGGGAGCTC 8760  
 CTTCCGTGGCCGACCCCGCTAACGATCTACTTGTGTAAGGCCACGGGCTCCGAG  
 8761 CTTGGGAACATCAGGCCCTGGGGCGCTTCCGGGCTGTCCGGCACCTCGACCCGAAA 8820  
 GAACGCCCTGTAGTCCGGGACCGCGCGAGAACCCCCGACAGGGCTGGAGCTGGCTT  
 8821 CCCGAGCGTTCCGACCCGTATCGACGACAAGACTGCCGGCGCGATTCCACCGCCGC 8880  
 GGGCTCCAAAGGGTGGCATAGCTGCTGTTGACGGCCCGCGTAAGGTGGCG  
 8881 CGCGGGCGGGCATCAGGACCGCAAGAACGGCTGCCCTTACTCGGCCACATGGGCAA 8940  
 GCGCCGCCGCCCCGTAGTCCTGGGTTCTCCGACGCCGAATGAGCCGGTGTACCCGTT  
 8941 GATAGGACTGCTGGCGCCGAGATCCCCGGCTGCAGGAATTGATATCAAGCTTATCG 9000  
 CTATCCGTGACGAGCCGGGCTTAGGGGCCCCGACGTCTTAAGCTATAGTCGAATAGC

Fig. 24/10

9001 ATACCGTCGACCTCGAGGGGGGCCGGTACCCAGCTTTGTTCCCTTAGTGAGGGTTA 9060  
 9001 -----+-----+-----+-----+-----+-----+-----+  
 9001 TATGGCAGCTGGAGCTCCCCCGGGCCATGGGTGAAACAGGGAAATCACTCCCAAT  
  
 9061 ATTGCGCGTTGGCGTAAATCATGGTCATAGCTGTTCTGTGTGAAATTGTTATCCGCTC 9120  
 9061 -----+-----+-----+-----+-----+-----+-----+  
 9061 TAACGCGCGAACCGCATTAGTACCAAGTATCGACAAGGGACACACTTTAACAAAGGCGAG  
  
 9121 ACAATTCCACACAACTACGAGCCGGAACGATAAGTGTAAAGCTGGGTGCCTAATGA 9180  
 9121 -----+-----+-----+-----+-----+-----+-----+  
 9121 TGTTAAAGGTGTGTTGTTATGCTCGGCTTCGTATTCACATTTGCGACCCACGGATTACT  
  
 9181 GTGAGCTAACTCACATTAATTGCGTTGCCTCACTGCCGCTTCCAGTCGGAAACCTG 9240  
 9181 -----+-----+-----+-----+-----+-----+-----+  
 9181 CACTCGATTGAGTGTAAATTACGCAACGCGAGTGACGGCGAAAGGTCAAGCCCTTGGAC  
  
 9241 TCGTGCCAGCTGCATTAATGAAATCGGCCAACGCGCGGGAGAGGCGGTTGCGTATTGGG 9300  
 9241 -----+-----+-----+-----+-----+-----+-----+  
 9241 AGCACGGTCGACGTAATTACTTAGCCGGTTGCGCCCTCTCCGCAAACGCAATAACCC  
  
 9301 CGCTCTCCGCTTCCTCGCTCACTGACTCGCTCGCTCGGTGTTGGCTGCGCGAGCG 9360  
 9301 -----+-----+-----+-----+-----+-----+-----+  
 9301 CGCAGAAGGCGAACGGAGCGAGTGACTGACCGACGGAGCCAGCAAGCCACGCCGCTCGC  
  
 9361 GTATCAGCTCACTCAAAGGCCTAATACGGTTATCCACAGAATCAGGGATAACGCAGGA 9420  
 9361 -----+-----+-----+-----+-----+-----+-----+  
 9361 CATACTCGAGTGAGTTCCGCCATTATGCCAATAGGTGTCTTAGTCCCCTATTGCGTCCT  
  
 9421 AAGAACATGTGAGCAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAGGCCCGTTGCTG 9480  
 9421 -----+-----+-----+-----+-----+-----+-----+  
 9421 TTCTTGTACACTCGTTTCCGGTCGTTTCCGGTCCTGGCATTTCCGGCGAACGAC  
  
 9481 CGGTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAG 9540  
 9481 -----+-----+-----+-----+-----+-----+-----+  
 9481 CGCAAAAAGGTATCCGAGGCCGGGGACTGCTCGTAGTGTAGCTGCGAGTCAGTC  
  
 9541 AGGTGGCAAACCCGACAGGACTATAAGATACCAGCGTTTCCCTGGAAAGCTCCCTC 9600  
 9541 -----+-----+-----+-----+-----+-----+-----+  
 9541 TCCACCGCTTGGCTGTCCTGATATTCTATGGTCGCAAAGGGGGACCTTCGAGGGAG  
  
 9601 GTGCGCTCTCTGTTCCGACCCCTGCCGTTACCGGATACCTGTCCGCCCTTCTCCCTCG 9660  
 9601 -----+-----+-----+-----+-----+-----+-----+  
 9601 CACCGGAGAGGACAAGGCTGGGACGGCGAATGGCTATGGACAGCCGGAAAGAGGGAAAG  
  
 9661 GGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTCGGTAGGTGCTT 9720  
 9661 -----+-----+-----+-----+-----+-----+-----+  
 9661 CCTTCGCACCGCGAAAGACTATCGAGTGCACATCCATAGAGTCAGGCCACATCCAGCAA  
  
 9721 CGCTCCAAGCTGGCTGTGACGAACCCCCGGTTAGCCGACCGCTGCCCTTATCC 9780  
 9721 -----+-----+-----+-----+-----+-----+-----+  
 9721 GCGAGGTTGCGACCCGACACACGTGCTGGGGCAAGTCGGGCTGGCGACGGGAATAGG  
  
 9781 GGTAACATCGTCTTGAGTCCAACCGGTAAGACACGACTATCGCACTGGCAGCAGCC 9840  
 9781 -----+-----+-----+-----+-----+-----+-----+  
 9781 CCATTGATAGCAGAACTCAGGTTGGGCATTCTGTGCTGAATAGCGGTGACCGTCGTCGG  
  
 9841 ACTGGTAACAGGATTAGCAGAGCGAGGTATGAGGGGTGCTACAGAGTTCTGAAGTGG 9900  
 9841 -----+-----+-----+-----+-----+-----+-----+  
 9841 TGACCAATTGTCCTAATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACCTCACC  
  
 9901 TGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGGCTCTGCTGAAGCCA 9960  
 9901 -----+-----+-----+-----+-----+-----+-----+  
 9901 ACCGGATTGATGCCGATGTGATCTTCTGTCAAAACCATAGACGCGAGACGACTTCGGT

Fig. 24/11

9961 GTTACCTTCGGAAAAAGAGTTGGTAGCTTGTATCCGGCRAACAAACACCACCGCTGGTAGC 10020  
 CAATGGAAGCCTTTCTAACCATCGAGAACTAGGCCCTTGTGGTGGCGACCATCG  
  
 10021 GGTGGTTTTTGTGCAAGCAGCAGATTACGCGCAGAAAAAGGGATCTCAAGAAGAT 10080  
 CCACCAAAAAACAAACGTTCGTCGTCTAATGCGCGCTTTCTAGAGTTCTCTA  
  
 10081 CCTTTGATCTTCTACGGGCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATT 10140  
 GGAAACTAGAAAAGATCCCCAGACTGCGAGTCACCTGCTTTGAGTGCACCTTCTAA  
  
 10141 TTGGTCATGAGATTATCAAAAAGGATCTCACCTAGATCCCTTTAAATTAAAAATGAAGT 10200  
 AACCAGTACTCTAATAGTTTCTAGAAGTGGATCTAGGAAATTAAATTTTACTTCA  
  
 10201 TTTAAATCAATCTAAAGTATATGAGTAAACTGGCTGACAGTACCAATGCTTAATC 10260  
 AAATTAGTTAGATTTCATATATACTCATTTGAACCGAGACTGTCAATGGTACGAATTAG  
  
 10261 AGTGAGGCACCTATCTCAGCGATCTGTCTATTGCTCATCCATAGTGCCTGACTCCCC 10320  
 TCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCAAGTAGGTATCACCGACTGAGGGG  
  
 10321 GTCGTGTAGATAACTACGATAACGGGAGGGCTTACCATCTGGCCCCAGTGCCTGCAATGATA 10380  
 CAGCACATCTATTGATGCTATGCCCTCCCGATGGTAGACCGGGGTACGACGTTACTAT  
  
 10381 CCGCGAGACCCACGCTCACCGCTCCAGATTATCAGCAATAACCAAGCCAGCCAGCGGAAGG 10440  
 GGCCTCTGGGTGCGAGTGGCCAGGTCTAAATAGTCGTTATTGGTCGGTCGGCTTCC  
  
 10441 GCGGAGCGCAGAAGTGGCTGCAACTTTATCCGCTCCATCCAGTCATTAATTGTC 10500  
 CGGCTCGCGTCTCACCGAGCGTTAAATAGCGGAGGTAGGTAGATAATTAAACACG  
  
 10501 CGGGAAAGCTAGAGTAAGTAGTCGCCAGTTAATAGTTGGCAACGTTGTCATTGCT 10560  
 GCGCTTCGATCTCATTCAAGCGTCATTATCAACCGGTTGCAACACGGTAACGA  
  
 10561 ACAGGGCATCGTGGTGTACGCTCGTCTGGTATGGCTTCAATTCACTCCGGTCCCAA 10620  
 TGTCGGTAGCACACAGTGCAGCAGCAACCATACCGAAGTAAGTCGAGGCCAGGGT  
  
 10621 CGATCAAGGCAGTTACATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCCCTCGGT 10680  
 GCTAGTTCCGCTCAATGTACTAGGGGTACAACACGTTTTCGCCATCGAGGAAGCCA  
  
 10681 CCTCCGATCGTTGTCAGAAGTAAGTGGCCAGTGTATCACTCATGGTTATGGCAGCA 10740  
 GGAGGGTAGCAACAGTCTCATTCAACCGGCGTCACAATAGTGAAGTACCAATACCGTCGT  
  
 10741 CTGCATAATTCTCTACTGTCATGCCATCCGTAAGATGCTTTCTGTGACTGGTAGTAC 10800  
 GACGTATTAAGAGAAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCACTCATG  
  
 10801 TCAACCAAGTCATTCTGAGAATAGTGTATGCCGACCCAGTTGCTCTTGCCCGCGTCA 10860  
 AGTTGGTTCAAGTAAACTCTTATCACATACGCCGCTGGCTAACGAGAACGGCCCGAGT  
  
 10861 ATACGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCCTCATCATTGGAAAACGT 10920  
 TATGCCCTATTATGGCCGGTGTATCGTCTGAAATTTCACGAGTAGTAACTTGTCA  
  
 10921 TCTCGGGCGAAAAGCTCAAGGATCTAACCGCTGTTGAGATCCAGTTGATGTAACCC 10980  
 AGAAGCCCCGCTTTGAGAGTCCCTAGAATGGCGACAACTCTAGGTCAAGCTACATTGGG

Fig. 24/12

10981 A C T C G T G C A C C C A A C T G A T C T T C A G G A T C T T T A C T T C A C C A G G G T T C T G G G T G A C C A 11040  
 TG A G G C A C G T G G G T T G A C T A G A A G T C G T A G A A A A T G A A A G T G G T C G C A A A G A C C C A C T C G T  
 A A A A C A G G A A G G C A A A A T G C C G C A A A A A G G G A A T A A G G G C G A C C G G A A A T G T T G A A T A  
 11041 T T T T G T C C T T C C G T T T T A C G G C G T T T T C C C T T A T T C C C G C T G T G C C T T T A C A A C T T A T 11100  
 C T C A T A C T C T T C C T T T T C A A T A T T A T T G A A G C A T T T A T C A G G G T T A T T G T C T C A T G A G C 11160  
 G A G T A T G A G A A G G A A A A A G T T A T A A T A A C T T C G T A A A T A G T C C C A A T A A C A G A G T A C T C G  
 G G A T A C A T A T T G A A T G T A T T T A G A A A A A A A C A A A T A G G G G T T C C G C G C A C A T T T C C C 11220  
 C C T A T G T A T A A A C T T A C A T A A A T C T T T T A T T G T T A T C C C C A A G G C G C G T G T A A A G G G  
 C G A A A A G T G C C A C  
 11221 G C T T T T C A C G G T G 11233

Digit. 25

Argala  
CGTGCT  
-----  
721 726  
CCACCA

Fig. 26

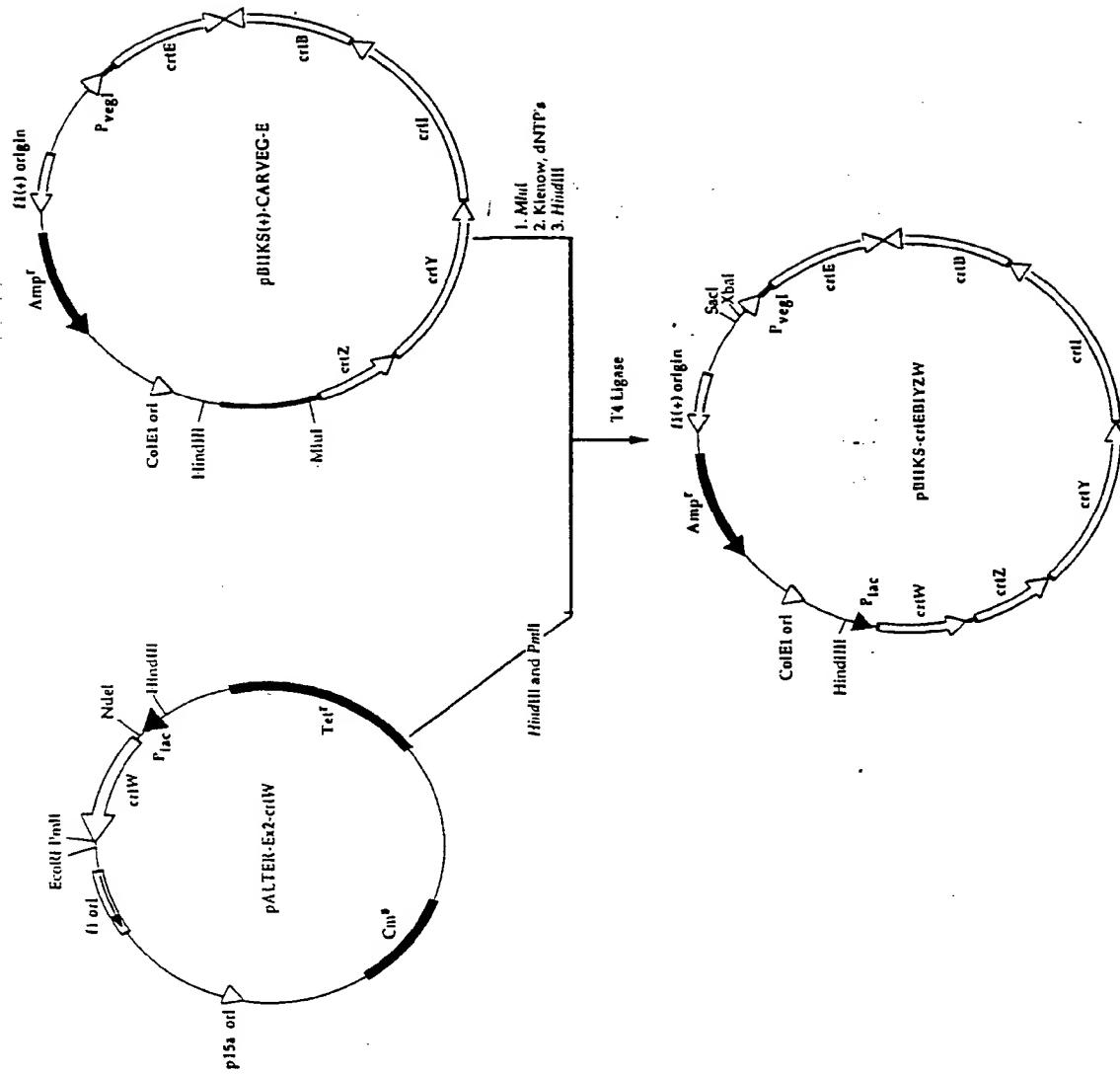


Fig. 27

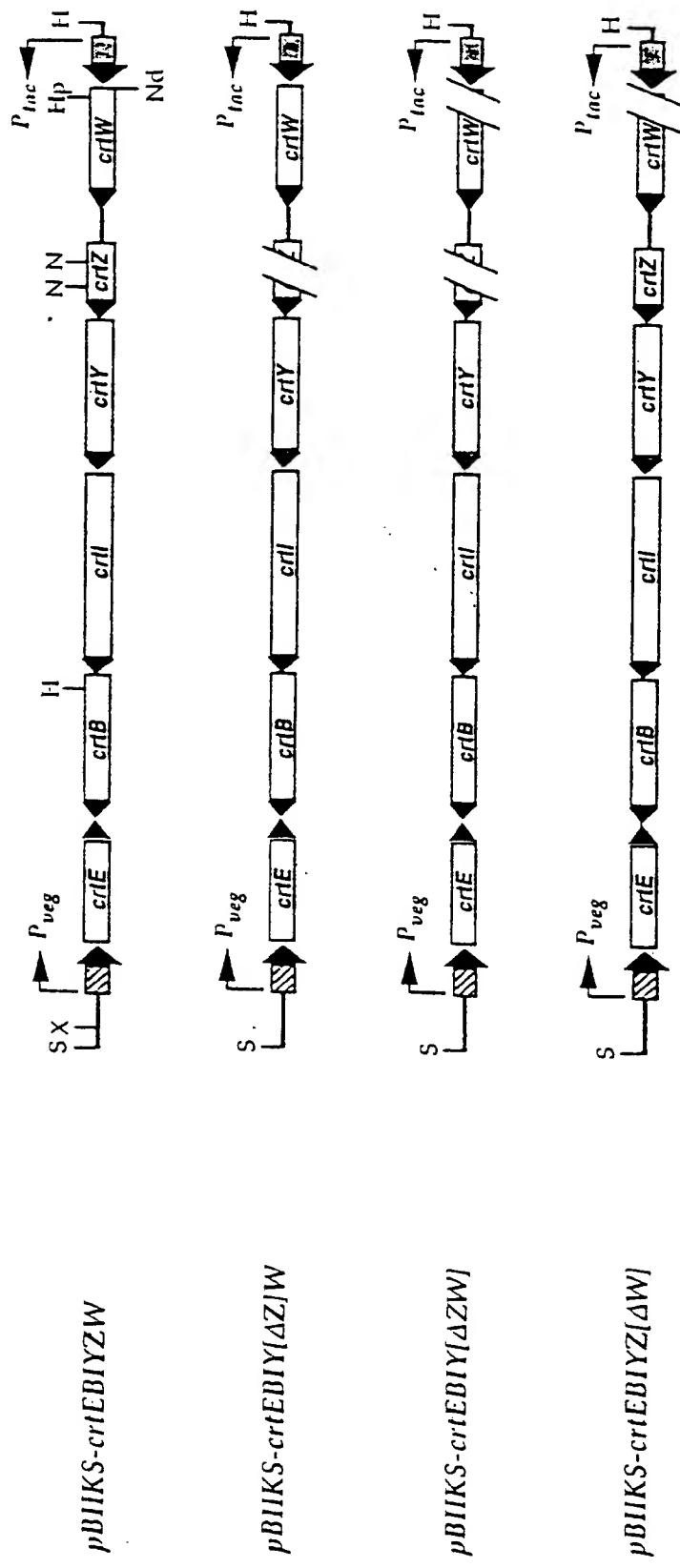
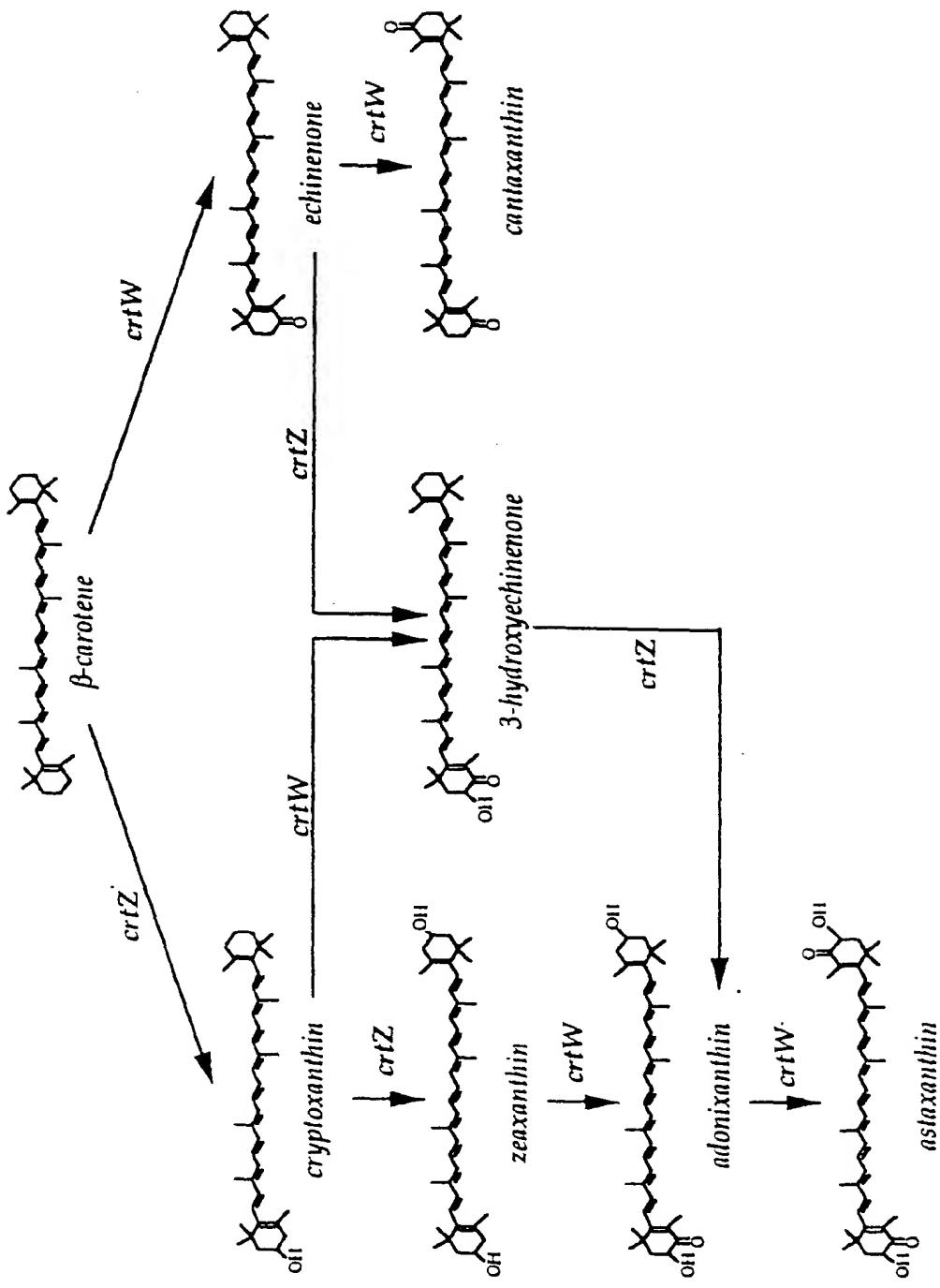


Fig. 28



(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 747 483 A3

(12)

## EUROPEAN PATENT APPLICATION

(88) Date of publication A3:  
07.05.1997 Bulletin 1997/19

(51) Int. Cl.<sup>6</sup>: C12N 15/53, C12N 15/54,  
C12N 15/60, C12N 15/70,  
C12N 15/74, C12N 15/80,  
C12N 15/81, C12P 23/00

(43) Date of publication A2:  
11.12.1996 Bulletin 1996/50

(21) Application number: 96108556.0

(22) Date of filing: 29.05.1996

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB IT LI NL

• Pasamontes, Luis  
4632 Trimbach (CH)  
• Tessier, Michel  
68100 Mulhouse (FR)  
• van Loon, Adolphus  
4310 Rheinfelden (CH)

(30) Priority: 09.06.1995 EP 95108888

(74) Representative: Braun, Axel et al  
F.Hoffmann-La Roche AG  
Patent Department (PLP),  
124 Grenzacherstrasse  
4070 Basel (CH)

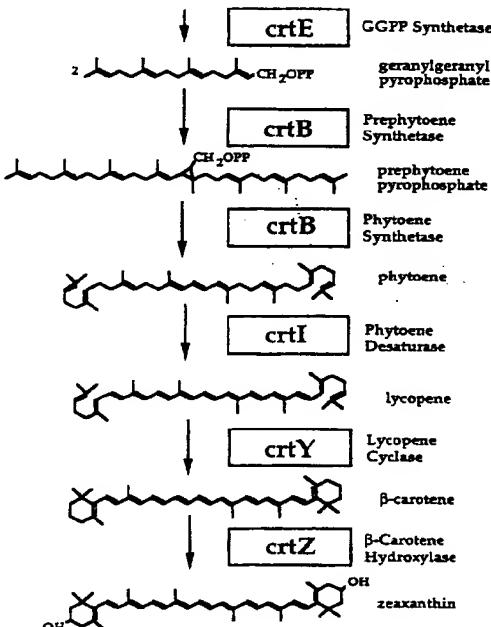
(71) Applicant: F. HOFFMANN-LA ROCHE AG  
4002 Basel (CH)

(72) Inventors:  
• Hohmann, Hans-Peter  
79100 Freiburg (DE)

### (54) Fermentative carotenoid production

(57) The present invention is directed to a DNA sequence comprising one or more DNA sequences selected from the group consisting of a DNA sequence which encodes the GGPP synthase of *Flavobacterium* sp. R1534 (crtE), a DNA sequence which encodes the prephytoene synthase of *Flavobacterium* sp. R1534 (crtB), a DNA sequence which encodes the phytoene desaturase of *Flavobacterium* sp. R1534 (crtI), a DNA sequence which encodes the lycopene cyclase of *Flavobacterium* sp. R1534 (crtY) or a DNA sequence which encodes the  $\beta$ -carotene hydroxylase of *Flavobacterium* sp. R1534 (crtZ) or DNA sequences which are substantially homologous, vectors comprising such DNA sequences and/or a DNA sequence which encodes the  $\beta$ -carotene  $\beta4$ -oxygenase of *Alcaligenes* strain PC-1 (crt W) or a DNA sequence which is substantially homologous, cells which are transformed by such DNA sequences and/or vectors, a process for the preparation of a desired carotenoid or a mixture of carotenoids by culturing such transformed cells and a process for the preparation of a food or feed composition.

Fig. 1



EP 0 747 483 A3



## EUROPEAN SEARCH REPORT

Application Number  
EP 96 10 8556

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.)
Y	US 3 891 504 A (A.J. SCHOCHER ET AL.) 24 June 1975 * the whole document * ---	1-35	C12N15/53 C12N15/54 C12N15/60 C12N15/70
D, Y	EP 0 393 690 A (KIRIN BEER KABUSHIKI KAISHA) 24 October 1990 * the whole document * ---	1-35	C12N15/74 C12N15/80 C12N15/81 C12P23/00
D, A	BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 209, no. 3, 26 April 1995, ORLANDO, FL US, pages 867-876, XP000602246 N. MISAWA ET AL.: "CANTHAXANTHIN BIOSYNTHESIS BY THE CONVERSION OF METHYLENE TO KETO GROUPS IN A HYDROCARBON BETA-CAROTENE BY A SINGLE GENE" * the whole document * ---	3,5,6, 18-20, 22,29-31	
D, A	WO 91 13078 A (AMOCO CORPORATION) 5 September 1991 * abstract * ---	1-35	TECHNICAL FIELDS SEARCHED (Int.Cl.)
A	DATABASE WPI Section Ch, Week 9428 Derwent Publications Ltd., London, GB; Class C03, AN 94-230229 XP002027030 & JP 06 165 683 A (KAIYO BIOTECHNOLOGY KENKYUSHO KK), 14 June 1994 * abstract * ---	27,34	
A	WO 91 03571 A (APPLIED FOOD BIOTECHNOLOGY INC) 21 March 1991 * page 1 - page 5 * ---	27,32	C12N C12P
The present search report has been drawn up for all claims			
Place of search EPO FORM 1500.02.82 (PNC01)	Date of completion of the search 6 March 1997	Examiner De Kok, A	
CATEGORY OF CITED DOCUMENTS		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ..... & : member of the same patent family, corresponding document	
X : particularly relevant if taken alone	Y : particularly relevant if combined with another document of the same category		
A : technological background	O : non-written disclosure		
P : intermediate document			



European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number

DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim
A	CH 595 444 A (SOCIÉTÉ DES PRODUITS NESTLÉ S.A.) 15 February 1978 * the whole document * -----	27,28
TECHNICAL FIELDS SEARCHED (Int.Cl.6)		
The present search report has been drawn up for all claims		
Place of search	Date of completion of the search	Examiner
BERLIN	6 March 1997	De Kok, A
CATEGORY OF CITED DOCUMENTS		
X : particularly relevant if taken alone	T : theory or principle underlying the invention	
Y : particularly relevant if combined with another document of the same category	E : earlier patent document, but published on, or after the filing date	
A : technological background	D : document cited in the application	
O : non-written disclosure	L : document cited for other reasons	
P : intermediate document	& : member of the same patent family, corresponding document	

